

# Explicit Reformulations for Robust Optimization Problems with General Uncertainty Sets

Averbakh, I; Zhao, Yun-Bin

DOI:  
[10.1137/060650003](https://doi.org/10.1137/060650003)

*Document Version*  
Peer reviewed version

*Citation for published version (Harvard):*  
Averbakh, I & Zhao, Y-B 2008, 'Explicit Reformulations for Robust Optimization Problems with General Uncertainty Sets', *SIAM Journal on Optimization*, vol. 18, no. 4, pp. 1436-1466.  
<https://doi.org/10.1137/060650003>

[Link to publication on Research at Birmingham portal](#)

## General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# EXPLICIT REFORMULATIONS FOR ROBUST OPTIMIZATION PROBLEMS WITH GENERAL UNCERTAINTY SETS

IGOR AVERBAKH\* AND YUN-BIN ZHAO†

(Published in *SIAM Journal on Optimization*, 18 (2007), No. 4, pp. 1436-1466)

**Abstract.** We consider a rather general class of mathematical programming problems with data uncertainty, where the uncertainty set is represented by a system of convex inequalities. We prove that the robust counterparts of this class of problems can be equivalently reformulated as finite and explicit optimization problems. Moreover, we develop simplified reformulations for problems with uncertainty sets defined by convex homogeneous functions. Our results provide a unified treatment of many situations that have been investigated in the literature, and are applicable to a wider range of problems and more complicated uncertainty sets than those considered before. The analysis in this paper makes it possible to use existing continuous optimization algorithms to solve more complicated robust optimization problems. The analysis also shows how the structure of the resulting reformulation of the robust counterpart depends both on the structure of the original nominal optimization problem and on the structure of the uncertainty set.

**Key words.** Robust optimization, data uncertainty, mathematical programming, homogeneous functions, convex analysis

**AMS subject classifications.** 90C30, 90C15, 90C34, 90C25, 90C05.

**1. Introduction.** In classical optimization models, the data are usually assumed to be known precisely. However, there are numerous situations where the data are inexact/uncertain. In many applications, the optimal solution of the nominal optimization problem may not be useful because it may be highly sensitive to small changes of the parameters of the problem.

Sensitivity analysis and stochastic programming are two traditional methods to deal with uncertain optimization problems. The former offers only local information near the nominal values of the data, while the latter requires one to make assumptions about the probability distribution of the uncertain data which may not be appropriate. Moreover, the stochastic programming approach often leads to very large optimization problems, and cannot guarantee satisfaction of certain hard constraints which is required in some practical settings.

An increasingly popular approach to optimization problems with data uncertainty is *robust optimization*, where it is assumed that possible values of data belong to some well-defined *uncertainty set*. In robust optimization, the goal is to find a solution that satisfies all constraints for any possible scenario from the uncertainty set, and optimizes the worst-case (guaranteed) value of the objective function. See e.g. [5]-[14], [21]-[26] and [29, 35, 39, 40]. The solutions of robust optimization models are “uniformly good” for realizations of data from the uncertainty set. Early work in this direction was done by Soyster [39, 40] and Falk [22] under the name of “inexact linear programming”. The robust optimization approach has been applied to various

---

\*Division of Management, University of Toronto at Scarborough, Scarborough, Ontario M1C 1A4, Canada (averbakh@utsc.utoronto.ca). The research of this author was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

†Institute of Applied Mathematics, AMSS, Chinese Academy of Sciences, Beijing 100080, China (ybzha@amss.ac.cn). Also with Division of Management, University of Toronto at Scarborough, Ontario M1C 1A4, Canada. The research of this author was supported by the Grant #10671199 and #70221001 from National Natural Science Foundation of China, and partially supported by CONACyT-SEP project #SEP-2004-C01-45786, Mexico.

problems in operations management, financial planning, and engineering design (e.g., [29, 26, 10, 6, 31, 35]).

A formulation of a robust model as a mathematical programming problem is called a *robust counterpart*. Since in the robust approach the constraints must be satisfied for all possible realizations of data from the uncertainty set, the robust counterpart is typically a complicated semi-infinite optimization problem. A fundamental question in robust optimization is whether the robust counterpart can be represented as a single finite and explicit optimization problem, so that existing optimization methods can be used to solve it. Such an analysis also helps to understand computational complexity of robust optimization problems.

So far, to obtain sufficiently simple robust counterparts, the uncertainty set was normally assumed to have a fairly simple structure, for example a Cartesian product of intervals, or an ellipsoid, or an intersection of ellipsoids, or a set defined by certain norms (see for example, [1]-[14], [23]-[26], [29]). Of course, the simpler the uncertainty set is, the easier it is to solve the robust optimization problem, and in some situations simplifying assumptions about uncertainty sets are natural when modelling a practical problem. However, more complicated uncertainty sets may be encountered in both theoretical study and in applications (see Remark 3.1 of this paper for details). Therefore, it is important to understand possibilities of the robust approach dealing with problems involving complicated or general uncertainty sets. Study of robust optimization problems with general uncertainty sets may provide additional tools for modelling intricate real-life situations and a unified treatment of specialized cases. Moreover, such a study can provide additional insights and results and even improve known results for some specialized cases when general results are reduced to such specialized cases (see Section 6 for details).

In this paper, we consider robust optimization problems with uncertainty sets defined by a system of convex inequalities. The optimization problems we consider may be non-convex and are wide enough to include linear programming, linear complementarity problems, quadratic programming, second order cone programming, and general polynomial programming problems. We prove that the robust counterparts of the considered problems with uncertainty are finite optimization problems which can be formulated by using the nominal data of the underlying optimization problem and the conjugates of the functions defining the uncertainty set. Compared with the original optimization problem, a major extra difficulty of the robust counterpart comes from the conjugates of the functions that define the uncertainty set. The conjugates of these functions usually are not given explicitly, and may be difficult to compute. To identify explicit and simplified formulations of robust counterparts, we focus on a class of convex functions whose conjugates can be expressed explicitly. Our strongest results and simplest reformulations of robust counterparts correspond to the case where the uncertainty sets are defined by convex homogeneous functions. This class of uncertainty sets is broad enough to include most uncertainty models that have been investigated in the literature, as well as many other important cases, for example where deviations of data from nominal values may be asymmetric and not even defined by norms.

We note that instead of optimizing the worst-case value of the objective function, another possibility is to optimize the worst-case regret, which is the worst-case deviation of the objective function value from the optimal value under the realized scenario, or, in other words, to minimize the worst-case loss in the objective function value that may occur because the decision is made before the realized scenario is known. This

criterion leads to minmax regret optimization models [29, 1, 2, 3, 4]. Minmax regret problems are typically computationally hard [29, 4], although there are exceptions (e.g., [1, 2, 3]). Minmax regret problems also fit the general paradigm of robust optimization, but we do not consider them in this paper. We note also that there are other concepts of robustness in the literature under the name of “model uncertainty” or “ambiguity”. See e.g. [42, 28, 17, 33, 18, 38, 27, 37, 16, 19, 20, 21, 31, 41].

This paper is organized as follows. In Section 2, we describe the class of optimization problems that we consider. In Section 3, we define the uncertainty set of data, and provide an equivalent, deterministic representation of the robust optimization problems via Fenchel’s conjugate functions. In Section 4, we give an explicit representation for the robust counterpart when the uncertainty set is defined by (non-homogeneous) convex functions that fall in the linear space generated by homogeneous functions of arbitrary degrees. The case of uncertainty sets defined by homogeneous functions is studied in Section 5. Specializing the general results of Sections 3, 4, and 5 to robust problems where the nominal problem is a linear programming problem and/or the uncertainty set is of a special type commonly used in the literature is discussed in Section 6, and concluding remarks are provided in Section 7.

**2. A class of optimization problems with data uncertainty.** We consider the following optimization problem:

$$(2.1) \quad \min\{c^T x : f_i(x) \leq b_i, i = 1, \dots, m, F(x) \leq 0\},$$

where  $c = (c_1, \dots, c_n)^T$  and  $b = (b_1, \dots, b_m)^T$  are fixed vectors, and  $f_i$ ’s are functions of the form

$$(2.2) \quad f_i(x) = \left(W^{(i)}(x)\right)^T M^{(i)} V^{(i)}(x), \quad i = 1, \dots, m,$$

where  $W^{(i)}(x)$  and  $V^{(i)}(x)$  are two mappings from  $R^n$  to  $R^{N_i}$ , and  $M^{(i)}$  is an  $N_i \times N_i$  real matrix,  $N_i$ ’s are positive integers. We write  $W^{(i)}(x)$  and  $V^{(i)}(x)$  as  $W^{(i)}(x) = (W_1^{(i)}(x), \dots, W_{N_i}^{(i)}(x))^T$  and  $V^{(i)}(x) = (V_1^{(i)}(x), \dots, V_{N_i}^{(i)}(x))^T$ , where each  $W_j^{(i)}$  ( $j = 1, \dots, N_i$ ) is a function from  $R^n$  to  $R$ .

We assume that only the data  $M^{(i)}$ ,  $i = 1, \dots, m$ , are subject to uncertainty. In (2.1),  $F(x) \leq 0$  denotes constraints without uncertainty, e.g. the simple constraints  $x \geq 0$ . We assume that  $c$  and  $b$  are certain without loss of generality, because a problem with uncertain  $c$  and  $b$  can be easily transformed into a problem with certain coefficients of the objective function and right-hand sides of the constraints. Also, if the objective function is not linear, it can be made linear by introducing an additional variable and a new constraint. We note that functions  $f_i$  are linear in the uncertain data  $M^{(i)}$  (but can be nonlinear in the decision variables  $x$ ).

The above optimization model is very general. For example, it includes the following important special cases.

*Linear Programming (LP).* Let  $A \in R^{m \times n}$  (i.e., an  $m \times n$  matrix) and  $b = (b_1, \dots, b_m)^T$ . Without loss of generality, we assume  $m \leq n$ . Consider functions  $f_i(x)$  of the form (2.2), where

$$W^{(i)}(x) = e_i \in R^n, \quad V^{(i)}(x) = x \in R^n, \quad M^{(i)} = \begin{bmatrix} A \\ 0 \end{bmatrix}_{n \times n},$$

where  $e_i$ , throughout this paper, denotes the  $i$ th column of  $n \times n$  Identity Matrix, and 0 in  $M^{(i)}$  denotes  $(n - m) \times n$  zero matrix. It is evident that the inequalities

$f_i = (W^{(i)})^T M^{(i)} V^{(i)} \leq b_i$ ,  $i = 1, \dots, m$ , are equivalent to  $Ax \leq b$ . Therefore, problem (2.1) with  $F(x) = -x \leq 0$  reduces to the linear programming problem:

$$(2.3) \quad \min\{c^T x : Ax \leq b, x \geq 0\}.$$

This implies that the linear programming problem (2.3) with uncertain coefficient matrix  $A$  is a special case of the optimization problem (2.1) with uncertain data  $M^{(i)}$ . There is also another way to write an LP in the form (2.1)-(2.2); see (6.14) and (6.15) in Section 6.2 for details.

*Linear Complementarity Problem (LCP).* Given a matrix  $M \in R^{n \times n}$  and a vector  $q \in R^n$ , the LCP is defined as

$$Mx + q \geq 0, \quad x \geq 0, \quad x^T(Mx + q) = 0.$$

Solutions to LCP are very sensitive to changes in data because of the equation  $x^T(Mx + q) = 0$ . When the matrix  $M$  is uncertain, it is hard to find a solution that satisfies the above system and is “immune” to changes of  $M$ . Thus, it is reasonable to consider the optimization form of LCP, i.e.,

$$\min\{x^T(Mx + q) : Mx + q \geq 0, x \geq 0\},$$

or equivalently

$$\min\{t : x^T(Mx + q) - t \leq 0, Mx + q \geq 0, x \geq 0\},$$

which is less sensitive in the sense that it is equivalent to LCP if the LCP has a solution, and can still have a solution even when the LCP has no solution. The above optimization problem can be reformulated as (2.2) by letting

$$W^{(1)}(x) = \begin{pmatrix} x \\ 1 \\ e_1 \end{pmatrix} \in R^{2n+1}; \quad W^{(i)} = \begin{pmatrix} 0^{(n+1)} \\ e_{i-1} \end{pmatrix} \in R^{2n+1}, \quad \text{for } i = 2, \dots, n+1,$$

$$M^{(i)} = \begin{bmatrix} M & q & 0 & \overbrace{0 \dots 0}^{n-1} \\ 0 & 0 & -1 & 0 \dots 0 \\ -M & -q & 0 & 0 \dots 0 \end{bmatrix}, \quad V^{(i)}(x) = \begin{pmatrix} x \\ 1 \\ t \\ 0^{(n-1)} \end{pmatrix} \in R^{2n+1}, \quad i = 1, \dots, n+1,$$

where  $t \in R$ , and  $0^{(n+1)}$  and  $0^{(n-1)}$  denote  $(n+1)$  and  $(n-1)$ -dimensional zero vectors, respectively. It is easy to verify that problem (2.1) with  $F(x) = -x \leq 0$  and  $f_i = (W^{(i)})^T M^{(i)} V^{(i)} \leq 0$  ( $i = 1, \dots, n+1$ ) is the same as the optimization form of LCP. It is worth mentioning that Zhang [43] considered equality constrained robust optimization, and his approach may be also used to deal with LCPs with uncertainty data.

*(Nonconvex) Quadratic Programming (QP).* Consider functions  $f_i(x)$  of the form (2.2) where

$$W^{(i)}(x) = \begin{pmatrix} x \\ 1 \end{pmatrix} \in R^{n+1}, \quad \text{for } i = 0, \dots, m,$$

$$V^{(0)}(x) = \begin{pmatrix} x \\ t \end{pmatrix} \in R^{n+1}, \quad V^{(i)}(x) = \begin{pmatrix} x \\ 0 \end{pmatrix} \in R^{n+1}, \quad \text{for } i = 1, \dots, m$$

and

$$(2.4) \quad M^{(i)} = \begin{bmatrix} Q_i & 0 \\ q_i^T & -1 \end{bmatrix}_{(n+1) \times (n+1)}, \quad \text{for } i = 0, \dots, m,$$

where each  $Q_i$  is an  $n \times n$  symmetric matrix and each  $q_i$  is a vector in  $R^n$ . Then the optimization problem (2.1) with the objective  $t$  and constraints  $f_i = (W^{(i)})^T M^{(i)} V^{(i)} \leq -c_i$  ( $i = 0, \dots, m$ ) is reduced to the quadratic programming problem:

$$\begin{aligned} \min \quad & x^T Q_0 x + q_0^T x + c_0 \\ \text{s.t.} \quad & x^T Q_i x + q_i^T x + c_i \leq 0, \quad \text{for } i = 1, \dots, m. \end{aligned}$$

Thus, a QP with uncertain coefficients  $(Q_i, q_i)$  ( $i = 0, \dots, m$ ) can be represented as an optimization problem (2.1) with uncertain data  $M^{(i)}$  given as (2.4).

*Second Order Cone Programming (SOCP).* Let  $A \in R^{m \times n}$ ,  $b \in R^m$ ,  $c \in R^n$ , and  $\beta$  be a scalar. Let

$$W^{(1)}(x) = V^{(1)}(x) = \begin{pmatrix} x \\ 1 \end{pmatrix} \in R^{n+1},$$

and

$$(2.5) \quad M^{(1)} = \begin{bmatrix} A^T A - cc^T & 0 \\ 2b^T A - 2\beta c^T & b^T b - \beta^2 \end{bmatrix},$$

and  $W^{(2)}(x) = e \in R^n$  (the vector with all components equal to 1),  $V^{(2)}(x) = x \in R^n$  and

$$M^{(2)} = \begin{bmatrix} -c^T \\ 0 \end{bmatrix}_{n \times n}.$$

Then the constraint  $f_1 = (W^{(1)})^T M^{(1)} V^{(1)} \leq 0$  together with  $f_2 = (W^{(2)})^T M^{(2)} V^{(2)} \leq \beta$  is equivalent to the second order cone constraint:  $\|Ax + b\| \leq c^T x + \beta$ . In fact,  $f_1 \leq 0$  can be written as

$$(Ax + b)^T (Ax + b) \leq (c^T x + \beta)^2$$

and  $f_2 \leq \beta$  can be written as  $c^T x + \beta \geq 0$ . Combination of these two inequalities leads to a second order cone constraint. Thus, uncertainty of the data  $(A, B, c, \beta)$  leads to uncertainty of the matrices  $M^{(1)}$  and  $M^{(2)}$ .

*Polynomial Programming.* We recall that a monomial in  $x_1, \dots, x_n$  is a product of the form  $x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ , where  $\alpha_1, \dots, \alpha_n$  are nonnegative integers. It is evident that if the components of  $W(x)$  and  $V(x)$  are monomials, then for any given matrix  $M$ , a function of the form (2.2) is a polynomial. Conversely, any real polynomial is a linear combination of some monomials, i.e.,

$$P(x_1, x_2, \dots, x_n) = \sum_{(\alpha_1, \alpha_2, \dots, \alpha_n)} C^{(\alpha_1, \alpha_2, \dots, \alpha_n)} x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

where  $C^{(\alpha_1, \dots, \alpha_n)}$  are real coefficients. Then the simplest way to write it in the form (2.2) is to set  $W(x) = e$ , set  $V(x)$  to be the vector of all monomials  $x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$  appearing in  $P(x)$ , and set  $M$  to be the diagonal matrix with diagonal entries  $C^{(\alpha_1, \alpha_2, \dots, \alpha_n)}$ . Thus polynomial optimization with uncertain coefficients is a special case of (2.1) with uncertain data  $M^{(i)}$ .

### 3. Robust counterparts as finite deterministic optimization problems.

We start with a description of the uncertainty set. Let  $K_i$ ,  $i = 1, \dots, m$ , be a bounded subset of  $R^{N_i^2}$  that contains the origin. Suppose that the uncertain data  $M^{(i)}$  ( $i = 1, \dots, m$ ) of the  $i$ th constraint of (2.1) are allowed to vary in such a way that the deviations from their fixed nominal values  $\overline{M}^{(i)}$  fall in  $K_i$ . That is, the uncertainty set of the data  $M^{(i)}$  is defined as

$$(3.1) \quad \mathcal{U}_i = \left\{ \widetilde{M}^{(i)} \mid \text{vec}(\widetilde{M}^{(i)}) - \text{vec}(\overline{M}^{(i)}) \in K_i \right\}, \quad i = 1, \dots, m,$$

where for a given matrix  $M$ ,  $\text{vec}(M)$  denotes the vector obtained by stacking the transposed rows of  $M$  on top of one another. Then the robust counterpart of the optimization problem (2.1) with uncertainty sets  $\mathcal{U}_i$  is defined as follows:

$$(3.2) \quad \min c^T x \\ \text{s.t. } f_i = \left( W^{(i)}(x) \right)^T \widetilde{M}^{(i)} V^{(i)}(x) \leq b_i, \forall \widetilde{M}^{(i)} \in \mathcal{U}_i, i = 1, \dots, m, F(x) \leq 0\},$$

which is a semi-infinite optimization problem. The optimal solution to this problem is feasible for all realizations of the data  $\widetilde{M}^{(i)}$ .

We denote by  $\delta(u|K)$  the indicator function of a set  $K$  (see [36]), and the conjugate function of  $\delta(u|K)$  is denoted by  $\delta^*(u|K)$  which is equal to the support function  $\psi_K(u) = \max\{u^T v : v \in K\}$ . First we state the following general result which shows that the robust counterpart (3.2) can be equivalently written as a finite deterministic optimization problem, regardless of the type of uncertainty sets.

**THEOREM 3.1.** *The robust optimization problem (3.2) is equivalent to the following finite and deterministic optimization problem:*

$$\begin{aligned} & \min c^T x \\ & \text{s.t. } \left( W^{(i)}(x) \right)^T \overline{M}^{(i)} V^{(i)}(x) + \delta^*(\chi_i | \text{cl}(\text{co}K_i)) \leq b_i, \quad i = 1, \dots, m, \\ & F(x) \leq 0, \end{aligned}$$

where  $\text{cl}(\text{co}K_i)$  denotes the closure of the convex hull of set  $K_i$ , and  $\chi_i = W^{(i)}(x) \otimes V^{(i)}(x) \in R^{N_i^2}$ , i.e., is the Kronecker Product of the vectors  $W^{(i)}(x)$  and  $V^{(i)}(x)$ .

*Proof.* In fact, the constraint  $f_i = \left( W^{(i)}(x) \right)^T \widetilde{M}^{(i)} V^{(i)}(x) \leq b_i$  for all  $\text{vec}(\widetilde{M}^{(i)}) - \text{vec}(\overline{M}^{(i)}) \in K_i$  is equivalent to

$$(3.3) \quad \sup \left\{ W^{(i)}(x)^T \widetilde{M}^{(i)} V^{(i)}(x) : \text{vec}(\widetilde{M}^{(i)}) - \text{vec}(\overline{M}^{(i)}) \in K_i \right\} \leq b_i.$$

Notice that for any square matrices  $B, C$ , we have  $\text{tr}(BC) = (\text{vec}(B))^T \text{vec}(C^T)$ . Thus, we have

$$\begin{aligned} \left( W^{(i)}(x) \right)^T \widetilde{M}^{(i)} V^{(i)}(x) &= \text{tr} \left( \widetilde{M}^{(i)} V^{(i)}(x) \left( W^{(i)}(x) \right)^T \right) \\ &= \left( \text{vec}(\widetilde{M}^{(i)}) \right)^T \text{vec} \left( W^{(i)}(x) \left( V^{(i)}(x) \right)^T \right) \\ &= \left( \text{vec}(\widetilde{M}^{(i)}) \right)^T \left( W^{(i)}(x) \otimes V^{(i)}(x) \right). \end{aligned}$$

Denoting  $\chi_i = W^{(i)}(x) \otimes V^{(i)}(x)$ , the constraint (3.3) can be written as

$$\begin{aligned} b_i &\geq \sup \left\{ \left( \text{vec}(\widetilde{M}^{(i)}) \right)^T \chi_i : \text{vec}(\widetilde{M}^{(i)}) - \text{vec}(\overline{M}^{(i)}) \in K_i \right\} \\ &= \left( \text{vec}(\overline{M}^{(i)}) \right)^T \chi_i + \sup_{u \in K_i} u^T \chi_i = \left( \text{vec}(\overline{M}^{(i)}) \right)^T \chi_i + \sup_{u \in \text{cl}(\text{co}K_i)} u^T \chi_i \\ &= \left( W^{(i)}(x) \right)^T \overline{M}^{(i)} V^{(i)}(x) + \delta^*(\chi_i | \text{cl}(\text{co}K_i)). \end{aligned}$$

The original semi-infinite constraints become finite and deterministic constraints.  $\square$

For robust optimization, when the uncertainty set is not convex, the robust counterpart remains unchanged if we replace the uncertainty set by its closed convex hull. This observation was first mentioned in [7], and can be seen clearly from the above result. Because of this fact, we may assume without loss of generality that each  $K_i$  is a closed convex set. In applications, the convex set  $K_i$  is usually determined by a system of convex inequalities. So, *throughout the rest of the paper, we assume that  $K_i$  is a closed, bounded convex set containing the origin and it can be represented as*

$$(3.4) \quad K_i = \left\{ u \mid g_j^{(i)}(u) \leq \Delta_j^{(i)}, j = 1, \dots, \ell^{(i)} \right\}, \quad i = 1, \dots, m,$$

where  $\ell^{(i)}$ 's are given integers,  $\Delta_j^{(i)}$ 's are constants, and  $g_j^{(i)}$ 's are proper closed convex functions from  $R^{N_i^2}$  to  $\overline{R}$ . Here,  $\overline{R} = R \cup \{+\infty\}$  and “proper” means that the function is finite somewhere (throughout the paper, we use the terminology from [36]). Since  $0 \in K_i$ , we have  $g_j^{(i)}(0) \leq \Delta_j^{(i)}$  for all  $j = 1, \dots, \ell^{(i)}$ .

*Remark 3.1.* In this remark, we give additional motivation for considering the general uncertainty set (3.4) as opposed to special uncertainty sets studied in the literature. We note that importance of studying robust problems with complicated uncertainty sets was emphasized, for example, in [15].

(i) Consider the following uncertainty set:

$$(3.5) \quad \mathcal{U} = \left\{ D \mid \exists z \in R^{|N|} : D = D_0 + \psi(z) = D_0 + \sum_{j \in N} \Delta D_j z_j, \|z\| \leq \Omega \right\},$$

where  $\Omega$  is a given number,  $D_0$  is a given vector (nominal values of the uncertain data), and  $\Delta D_j$ 's are directions of data perturbation. This uncertainty set has been widely used in the literature (e.g. [5]-[14], [23]-[26]). It is the image of a ball (defined by some norm) under linear transformation, i.e., the function  $\psi(z)$  here is a linear function in  $z$ . This widely used uncertainty set can be written in the form (3.4) with only one convex inequality  $g(u) \leq \Omega$ , where function  $g(u)$  is also homogeneous of 1-degree, and  $g(u)$  is not a norm in general unless  $|N|$  is equal to the number of data and the data perturbation directions  $\Delta D_j$ 's are linearly independent (see Section 6.1 for details). This typical example shows that it is necessary to study the case when the functions  $g_j^{(i)}(u)$  in (3.4) are convex and homogeneous (but not necessarily norms). Section 5 of this paper is devoted to this important case.

For the uncertainty set  $\mathcal{U}$  defined by (3.5), the function  $\psi(z)$  is linear in  $z$ . In some applications, however, such a model is insufficient for description of more complicated uncertainty sets. The next two examples show that in some situations the function  $\psi(z)$  may be nonlinear and hence the uncertainty set may be much more complicated.



(ii) Consider the second order cone programming (SOCP). It is often assumed that the data  $(A, b, c)$  are subject to an ellipsoidal uncertainty set which is the case of (3.5) where the norm is the 2-norm. When we reformulate SOCP into the form of (2.1), the data  $M^{(1)}$  is determined by the matrix (2.5). It is easy to see the data  $M^{(1)}$  belongs to the following uncertainty set

$$(3.6) \quad \mathcal{U} = \left\{ D \mid \exists z \in R^{|N|} : D = D^0 + \psi(z), \quad \|z\| \leq \Omega \right\},$$

where  $\psi(z)$  is a quadratic function in  $z$ . Thus, this example shows that a more complicated uncertainty set than (3.5) might appear when we make a reformulation of the problem. Such reformulations are often made when a problem is studied from different perspectives.

(iii) This example, taken from [23], shows that a nonlinear function  $\psi(z)$  arises in (3.6) when robust interpolation problems are considered. Let  $n \geq 1$  and  $k$  be given integers. We want to find a polynomial of degree  $n-1$ ,  $p(t) = x_1 + \dots + x_n t^{n-1}$  that interpolates given points  $(a_i, b_i)$ , i.e.,  $p(a_i) = b_i, i = 1, \dots, k$ . If interpolation points  $(a_i, b_i)$  are known precisely, we obtain the following linear equation

$$\begin{bmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_k & \cdots & a_k^{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Now assume that  $a_i$ 's are not known precisely, i.e.,  $a_i(\delta) = a_i + \delta_i, i = 1, \dots, k$ , where the  $\delta = (\delta_1, \dots, \delta_k)$  is unknown but bounded, i.e.,  $\|\delta\|_\infty \leq \rho$  where  $\rho \geq 0$  is given. A robust interpolant is a solution  $x$  that minimizes  $\|A(\delta)x - b\|$  over the region  $\|\delta\|_\infty \leq \rho$ , where

$$A(\delta) = \begin{bmatrix} 1 & a_1(\delta) & \cdots & a_1(\delta)^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_k(\delta) & \cdots & a_k(\delta)^{n-1} \end{bmatrix}$$

is an uncertain Vandermonde matrix. Such a matrix can be written in the form (3.6) with nonlinear function  $\psi(z)$ . In fact, we have (see [23] for details)

$$A(\delta) = A(0) + L\Delta(I - D\Delta)^{-1}R_A$$

where  $L, D$  and  $R_A$  are constant matrices determined by  $a_i$ 's, and  $\Delta = \oplus_{i=1}^k \delta_i I_{n-1}$ .

(iv) Our model provides a unified treatment of many uncertainty sets in the literature. Note that (3.6) can be written in the form (3.4), by letting  $g(D) = \inf\{\|z\| : D = \psi(z)\}$ . Then  $\mathcal{U} - \{D_0\} = \{D : g(D) \leq \Omega\}$ . This can be proved by the same argument as Lemma 6.1 in this paper.

(v) Studying problems with general uncertainty sets may in fact lead to new or stronger results for important special cases, as we demonstrate in Section 6.

Since robust optimization problems in general are semi-infinite optimization problems which are hard to solve, the fundamental question is whether a robust optimization problem can be explicitly represented as an equivalent finite optimization problem, so that the existing optimization methods can be applied. We are addressing this question in this paper. It should be mentioned that, generally, two research directions are possible: 1) Developing computationally tractable approximate (relaxed) formulations; 2) Developing exact formulations which, naturally, will be computationally difficult for sufficiently complicated nominal problems and/or uncertainty sets.

Our paper focuses on the second direction; the first direction was investigated, for instance, in Bertsimas and Sim [14]. We believe that both directions are important for theoretical and practical progress in robust optimization; we comment on this in more detail in Section 6.

Let us mention some auxiliary results and definitions. Given a function  $f$ , we denote its domain by  $\text{dom}(f)$ , and denote its Fenchel's conjugate function by  $f^*$ , i.e.,

$$f^*(w) = \sup_{x \in \text{dom}(f)} (w^T x - f(x)).$$

We recall that the infimal convolution function of  $g_j (j = 1, \dots, \ell)$ , denoted by  $g_1 \diamond g_2 \diamond \dots \diamond g_\ell$ , is defined as

$$(g_1 \diamond g_2 \diamond \dots \diamond g_\ell)(u) = \inf \left\{ \sum_{j=1}^{\ell} g_j(u_j) : \sum_{j=1}^{\ell} u_j = u \right\}.$$

The following result will be used in our later analysis.

LEMMA 3.2. ([36], Theorem 16.4.) *Let  $f_1, \dots, f_\ell : R^n \rightarrow \bar{R}$  be proper convex functions. Then  $(cl(f_1) + \dots + cl(f_\ell))^* = cl(f_1^* \diamond \dots \diamond f_\ell^*)$ , where  $cl(f)$  denotes the closure of the convex function  $f$ . If the relative interiors of the domains of these functions, i.e.,  $\text{ri}(\text{dom}(f_i))$ ,  $i = 1, \dots, \ell$ , have a point in common, then*

$$\left( \sum_{i=1}^{\ell} f_i \right)^* (x) = (f_1^* \diamond \dots \diamond f_\ell^*)(x) = \inf \left\{ \sum_{i=1}^{\ell} f_i^*(x_i) : \sum_{i=1}^{\ell} x_i = x \right\},$$

where for each  $x \in R^n$  the infimum is attained.

Now we consider the robust programming problem (3.2) where the uncertainty set is determined by (3.1) and (3.4). We have the following general result.

THEOREM 3.3. *Let  $K_i$  ( $i = 1, \dots, m$ ) be given by (3.4) where each  $g_j^{(i)}$  ( $j = 1, \dots, \ell^{(i)}$ ) is a closed proper convex function. Suppose that Slater's condition holds for each  $i$ , i.e., for each  $i$ , there exists a point  $u_0^{(i)}$  such that  $g_j^{(i)}(u_0^{(i)}) < \Delta_j^{(i)}$  for all  $j = 1, \dots, \ell^{(i)}$ . Then the robust counterpart (3.2) is equivalent to*

$$\begin{aligned} & \min c^T x \\ & \text{s.t. } \left( W^{(i)}(x) \right)^T \bar{M}^{(i)} V^{(i)}(x) + \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} + \left( \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} g_j^{(i)} \right)^* (\chi_i) \leq b_i, \quad i = 1, \dots, m, \\ & \lambda_j^{(i)} \geq 0, \quad j = 1, \dots, \ell^{(i)}; \quad i = 1, \dots, m, \\ & F(x) \leq 0, \end{aligned}$$

where  $\chi_i = W^{(i)}(x) \otimes V^{(i)}(x)$ . This problem can be further written as

$$\begin{aligned} & \min c^T x \\ & \text{s.t. } \left( W^{(i)}(x) \right)^T \bar{M}^{(i)} V^{(i)}(x) + \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} + \Upsilon^{(i)}(\lambda^{(i)}, u^{(i)}) \leq b_i, \quad i = 1, \dots, m, \\ (3.7) \quad & \chi_i = \begin{cases} \sum_{j \in J_i} u_j^{(i)}, & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases} \quad i = 1, \dots, m, \\ & \lambda_j^{(i)} \geq 0, \quad j = 1, \dots, \ell^{(i)}; \quad i = 1, \dots, m, \\ & F(x) \leq 0. \end{aligned}$$

where  $J_i = \{j : \lambda_j^{(i)} > 0, j = 1, \dots, \ell^{(i)}\}$ ,  $\lambda^{(i)}$  denotes the vector whose components are  $\lambda_j^{(i)}$ ,  $j = 1, \dots, \ell^{(i)}$ ,  $u^{(i)}$  denotes the vector whose components are  $u_j^{(i)}$ ,  $j \in J_i$ , and

$$\Upsilon^{(i)}(\lambda^{(i)}, u^{(i)}) = \begin{cases} \sum_{j \in J_i} \lambda_j^{(i)} \left(g_j^{(i)}\right)^* \left(u_j^{(i)} / \lambda_j^{(i)}\right), & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases}$$

*Proof.* We see from the proof of Theorem 3.1 that  $x$  is feasible to the robust problem (3.2) if and only if  $F(x) \leq 0$  and for each  $i$  we have

$$(3.8) \quad \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \max_{u \in K_i} u^T \chi_i \leq b_i.$$

Let  $Z(\chi_i) = \max\{u^T \chi_i : u \in K_i\}$  where  $K_i$  is given by (3.4) which by our assumption is a bounded, closed convex set. Thus the maximum value of the convex optimization problem  $\max\{u^T \chi_i : u \in K_i\}$  is finite and attainable. Denote the Lagrangian multiplier vector for this problem by  $\lambda^{(i)} = (\lambda_1^{(i)}, \lambda_2^{(i)}, \dots, \lambda_{\ell^{(i)}}^{(i)}) \in R_+^{\ell^{(i)}}$ . Since Slater's condition holds for the problem  $\max\{u^T \chi_i : u \in K_i\}$ , by Lagrangian Saddle-Point Theorem (see e.g. Theorem 28.3, Corollary 28.3.1 and Theorem 28.4 in [36]), we have

$$\begin{aligned} Z(\chi_i) &= -\min\{-u^T \chi_i : g_j^{(i)}(u) \leq \Delta_j^{(i)}, j = 1, \dots, \ell^{(i)}\} \\ &= -\sup_{\lambda^{(i)} \in R_+^{\ell^{(i)}}} \inf_{u \in R^{N_i^2}} \left(-u^T \chi_i + \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \left(g_j^{(i)}(u) - \Delta_j^{(i)}\right)\right) \\ &= -\sup_{\lambda^{(i)} \in R_+^{\ell^{(i)}}} \left[-\sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} + \inf_{u \in R^{N_i^2}} \left(-u^T \chi_i + \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} g_j^{(i)}(u)\right)\right] \\ &= -\sup_{\lambda^{(i)} \in R_+^{\ell^{(i)}}} \left[-\sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} - \sup_{u \in R^{N_i^2}} \left(u^T \chi_i - \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} g_j^{(i)}(u)\right)\right] \\ &= -\sup_{\lambda^{(i)} \in R_+^{\ell^{(i)}}} \left(-\sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} - \left(\sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} g_j^{(i)}\right)^* (\chi_i)\right) \\ (3.9) \quad &= \inf_{\lambda^{(i)} \in R_+^{\ell^{(i)}}} \left(\sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} + \left(\sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} g_j^{(i)}\right)^* (\chi_i)\right). \end{aligned}$$

Under our assumptions, the above infimum is attainable (by the existence of a saddle point of the Lagrangian function [36]). Substituting (3.9) into (3.8), we see that  $x$  satisfies (3.8) if and only if it satisfies the following inequalities for some  $\lambda^{(i)}$ :

$$(3.10) \quad \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} + \left(\sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} g_j^{(i)}\right)^* (\chi_i) \leq b_i,$$

$$(3.11) \quad \lambda^{(i)} = (\lambda_1^{(i)}, \lambda_2^{(i)}, \dots, \lambda_{\ell^{(i)}}^{(i)}) \in R_+^{\ell^{(i)}}.$$

Indeed, if  $x$  is feasible to (3.8), since the infimum in (3.9) is attainable, there exists some  $\lambda^{(i)} \in R_+^{\ell^{(i)}}$  such that  $(x, \lambda^{(i)})$  is feasible to the system (3.10)-(3.11). Conversely,

if  $(x, \lambda^{(i)})$  is feasible to (3.10) and (3.11), then by (3.9), we see that (3.10) implies (3.8). Replacing (3.8) by (3.10) together with (3.11), the first part of the desired result follows from Theorem 3.1.

We now derive the optimization problem (3.7). Suppose that  $(x, \lambda^{(i)})$  satisfies (3.10) and (3.11). We have two cases:

Case 1.  $J_i = \{j : \lambda_j^{(i)} > 0, j = 1, \dots, \ell^{(i)}\} \neq \emptyset$ . Denote by  $u^{(i)}$  the vector whose components are  $u_j^{(i)}, j \in J_i$ . Notice that for any constant  $\alpha > 0$ , the conjugate  $(\alpha f)^*(x) = \alpha f^*(x/\alpha)$ . For given  $\lambda^{(i)} \in R_+^{\ell^{(i)}}$ , by Lemma 3.2, we have

$$\left( \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} g_j^{(i)} \right)^* (\chi_i) = \inf_{u^{(i)}} \left\{ \sum_{j \in J_i} \lambda_j^{(i)} (g_j^{(i)})^* (u_j^{(i)} / \lambda_j^{(i)}) : \chi_i = \sum_{j \in J_i} u_j^{(i)} \right\}.$$

Again, by Lemma 3.2, the infimum above is attainable and hence there are  $u_j^{(i)}, j \in J_i$  such that

$$\begin{aligned} \left( \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} g_j^{(i)} \right)^* (\chi_i) &= \sum_{j \in J_i} \lambda_j^{(i)} (g_j^{(i)})^* (u_j^{(i)} / \lambda_j^{(i)}), \\ \chi_i &= \sum_{j \in J_i} u_j^{(i)}. \end{aligned}$$

Case 2.  $J_i = \emptyset$ . Notice that

$$\left( \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} g_j^{(i)} \right)^* (w) = \sup_{u \in R^n} (w^T u - 0) = \begin{cases} \infty, & \text{if } w \neq 0, \\ 0, & \text{if } w = 0. \end{cases}$$

Since  $(x, \lambda^{(i)})$  is feasible to (3.10) and (3.11), we conclude that for this case

$$\chi_i = 0, \quad \left( \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} g_j^{(i)} \right)^* (\chi_i) = 0.$$

Combining the above two cases leads to the optimization problem (3.7).  $\square$

We see from Theorem 3.3 that the level of complexity of the robust counterpart, compared with the nominal optimization problem, is determined mainly by the conjugate functions  $(g_j^{(i)})^* (j = 1, \dots, \ell^{(i)}, i = 1, \dots, m)$  and functions  $\chi_i (i = 1, \dots, m)$ . The more complicated the conjugate functions are, the more difficult the robust counterpart is. Notice that the constraint  $\sum_{j \in J_i} u_j^{(i)} = \chi_i$  is an explicit expression, and in some cases, e.g. LP,  $\chi_i$  is linear in  $x$ , and thus does not add difficulty. We also note that when  $\ell^{(i)} = 1$ , i.e., when  $K_i$  is defined by only one constraint, then  $u_j^{(i)} = \chi_i$ , in which case the formula  $\sum_{j \in J_i} u_j^{(i)} = \chi_i$  will not appear in (3.7). For an arbitrary function, however, its conjugate function is not given explicitly and hence (3.7) is not an explicit optimization problem. As a result, to obtain an explicit formulation of the robust counterpart, one has to compute the conjugate functions of the constraint functions  $g_j^{(i)}$ , which except for very simple cases is not easy. This motivates us to investigate in the remainder of the paper under what conditions the robust counterpart in Theorem 3.3 can be further simplified, avoiding the computation of conjugate functions.

**4. Explicit reformulation for robust counterparts.** For any function  $f$ , let

$$\mathcal{RD}(f) = \bigcup_{x \in \text{dom}(f)} \partial f(x),$$

that is,  $\mathcal{RD}(f)$  is the range of the subdifferential mapping  $\partial f(\cdot)$ . If  $f$  is differentiable,  $\mathcal{RD}(f)$  reduces to the range of its gradient mapping, i.e.,  $\mathcal{RD}(f) = \{\nabla f(x) : x \in \text{dom}(f)\}$ . In this section we make the following assumption.

**ASSUMPTION 4.1.** *The functions  $g_j^{(i)}(j = 1, \dots, \ell^{(i)}, i = 1, \dots, m)$  in (3.4) belong to the set of convex functions  $f$  that satisfy the condition*

$$(4.1) \quad \text{dom}(f^*) = \mathcal{RD}(f).$$

In fact, by the definition of subdifferential, the following relation always holds for any proper convex function:  $\text{dom}(f^*) \supseteq \mathcal{RD}(f)$ . Condition (4.1) requires the converse also to be true. Indeed, condition (4.1) holds for many functions. It is evident that all convex functions defined on a subset of  $R^n$  with  $\mathcal{RD}(f) = R^n$  satisfy condition (4.1). For example, when the function  $f$  is differentiable and strongly convex on  $R^n$ , the gradient  $\nabla f(x)$  is a strongly monotone function from  $R^n$  to  $R^n$ . This implies that  $\nabla f(x)$  is a bijective mapping ([34], Theorem 6.4.4), and hence we have  $\mathcal{RD}(f) = R^n$ . A simple example is the quadratic function  $f = \frac{1}{2}x^T Qx + bx + c$  where  $Q$  is a positive definite matrix, then  $\mathcal{RD}(f) = \{Qx + b : x \in R^n\} = R^n$ . When  $\mathcal{RD}(f) \neq R^n$ , (4.1) can still be satisfied in many cases. Later, we will show that all convex homogeneous of 1-degree functions satisfy (4.1) trivially, and  $\mathcal{RD}(f)$  of any function of this class is a closed bounded region including the origin. Notice that for any  $(u, x)$  such that  $u \in \partial f(x)$ , we have  $f^*(u) = u^T x - f(x)$ . The importance of condition (4.1) is that under (4.1), for any  $u \in \text{dom}(f^*)$  there is  $x \in \text{dom}(f)$  such that  $u \in \partial f(x)$  and therefore  $f^*(u) = u^T x - f(x)$ . Therefore, under Assumption 4.1, the robust counterpart (3.7) can be represented explicitly. However, we omit the statement of this general result. We are interested now in functions that have more properties leading to further simplification of the robust counterpart.

We recall that a function  $h : R^n \rightarrow \bar{R}$  is said to be positively homogeneous if there exists a constant  $p > 0$  such that  $h(\lambda x) = \lambda^p h(x)$  for all  $\lambda \geq 0$  and  $x \in \text{dom}(h)$ . If such a  $p$  exists, we simply say that the function  $h$  is homogeneous of  $p$ -degree. Notice that the definition implies  $0 \in \text{dom}(h)$  and  $h(0) = 0$ . We consider the linear space  $\mathcal{L}_H$  generated by homogeneous functions, i.e.,  $\mathcal{L}_H$  is the collection of all functions that are finite linear combinations of homogeneous functions. Notice that for any real number  $\alpha$ ,  $(\alpha h)(x)$  is also a homogeneous function if  $h$  is homogeneous. Therefore,  $\mathcal{L}_H$  is the set of all finite sums of homogeneous functions. Clearly, a function  $f$  which is the sum of several homogeneous functions  $f_i$  is not necessarily homogeneous, unless all  $f_i$  have the same homogeneous degree. Linear space  $\mathcal{L}_H$  includes many important classes of functions. Needless to say, all homogeneous functions (in particular, all norms  $\|\cdot\|$ ) are in  $\mathcal{L}_H$  and all polynomial functions are in  $\mathcal{L}_H$ .

The classical Euler's Homogeneous Function Theorem claims that if  $f$  is continuously differentiable and homogeneous of  $p$ -degree, then  $pf(x) = x^T \nabla f(x)$ , where  $\nabla f(x)$  is the gradient of  $f$ . Below, we establish a somewhat different version of the Euler's Homogeneous Function Theorem. This version allows the function to be non-differentiable and non-homogeneous, but belong to  $\mathcal{L}_H$  and be convex.

**LEMMA 4.1.** *Let  $f : R^n \rightarrow \bar{R}$  be a convex function in  $\mathcal{L}_H$ . Thus,  $f$  can be*

represented as  $f(x) = f_1(x) + \dots + f_N(x)$  for some  $N$ , where each  $f_i$  is homogeneous of  $p_i$ -degree, respectively.

(i) For any  $x \in \text{dom}(f)$ , we have

$$\sum_{i=1}^N p_i f_i(x) = \inf_{y \in \partial f(x)} y^T x = \sup_{y \in \partial f(x)} y^T x,$$

i.e., for any  $y \in \partial f(x)$ , we have  $\sum_{i=1}^N p_i f_i(x) = y^T x$ .

(ii) Suppose that  $f : R^n \rightarrow \bar{R}$  is a convex function and is homogeneous of  $p$ -degree. Then for any  $x \in \text{dom}(f)$ , and for any  $y \in \partial f(x)$ , we have  $pf(x) = y^T x$ .

*Proof.* For any given  $x \in \text{dom}(f)$  and  $y \in \partial f(x)$ , by definition of subdifferential we have  $f(u) \geq f(x) + y^T(u - x)$  for all  $u \in \text{dom}(f)$ . Notice that  $x \in \text{dom}(f)$  if and only if  $x \in \text{dom}(f_i)$  for all  $i = 1, \dots, N$ . Since all  $f_i$ 's are homogeneous, for any  $t > 0$ , we have  $u = tx \in \text{dom}(f_i)$  for all  $i = 1, \dots, N$ . This in turn implies that  $u = tx \in \text{dom}(f)$  for any  $t > 0$ . Setting  $u = tx$  in the above inequality and by using homogeneity, we have

$$f(tx) = \sum_{i=1}^N f_i(tx) = \sum_{i=1}^N t^{p_i} f_i(x) \geq f(x) + y^T(tx - x), \quad \text{for all } t > 0.$$

i.e.,

$$(4.2) \quad \sum_{i=1}^N (t^{p_i} - 1) f_i(x) \geq (t - 1) y^T x, \quad \text{for all } t > 0.$$

For  $t > 1$ , dividing both sides by  $t - 1$  and noting that  $y$  is any given element in  $\partial f(x)$ , we see from the above inequality that

$$\lim_{t \rightarrow 1+} \sum_{i=1}^N \frac{t^{p_i} - 1}{t - 1} f_i(x) \geq \sup_{y \in \partial f(x)} y^T x.$$

Thus, we have  $\sum_{i=1}^N p_i f_i(x) \geq \sup_{y \in \partial f(x)} y^T x$ . Similarly, when  $t < 1$ , dividing both sides of (4.2) by  $t - 1$ , we can prove that

$$\sum_{i=1}^N p_i f_i(x) = \lim_{t \rightarrow 1-} \sum_{i=1}^N \frac{t^{p_i} - 1}{t - 1} f_i(x) \leq \inf_{y \in \partial f(x)} y^T x.$$

Combining the last two inequalities yields the desired result (i). Setting  $N = 1$ , we obtain the result (ii) from (i).  $\square$

Notice that when  $N > 1$  Lemma 4.1 requires convexity of  $f$ , but does not require convexity of individual functions  $f_i$ , which can be nonconvex. The next theorem is the main result of this section, which states that the robust counterpart can be represented explicitly by using only the nominal data and the constraint functions  $g_i$  together with their sub-differentials.

**THEOREM 4.2.** Let  $K_i$  ( $i = 1, \dots, m$ ) be given by (3.4) where each  $g_j^{(i)}(j = 1, \dots, \ell^{(i)}, i = 1, \dots, m)$  is a closed proper convex function and belongs to the linear space  $\mathcal{L}_H$ , and is represented as

$$(4.3) \quad g_j^{(i)}(x) = \sum_{k=1}^{m^{(ij)}} h_k^{(ij)}(x),$$

where each  $h_k^{(ij)}(x)$  is homogeneous of  $p_k^{(ij)}$ -degree, and each  $m^{(ij)} \geq 1$  is a given integer number. Let  $g_j^{(i)}$  satisfy Assumption 4.1 and Slater's condition for each  $i$ . Then the robust programming problem (3.2) is equivalent to

$$\begin{aligned}
(4.4) \quad & \min c^T x \\
& \text{s.t. } \left( W^{(i)}(x) \right)^T \overline{M}^{(i)} V^{(i)}(x) + \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} + \Upsilon^{(i)} \leq b_i, \quad i = 1, \dots, m \\
& \chi_i = \begin{cases} \sum_{j \in J_i} u_j^{(i)}, & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases} \quad i = 1, \dots, m, \\
& \lambda_j^{(i)} \geq 0, \quad j = 1, \dots, \ell^{(i)}; \quad i = 1, \dots, m, \\
& F(x) \leq 0,
\end{aligned}$$

where  $\chi_i = W^{(i)}(x) \otimes V^{(i)}(x)$  and  $J_i = \{j : \lambda_j^{(i)} > 0, j = 1, \dots, \ell^{(i)}\}$ , and

$$\Upsilon^{(i)} = \begin{cases} \sum_{j \in J_i} \lambda_j^{(i)} \left( \sum_{k=1}^{m^{(ij)}} (p_k^{(ij)} - 1) h_k^{(ij)}(w_j^{(i)}) \right), & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise} \end{cases},$$

where  $w_j^{(i)}$  satisfies that  $u_j^{(i)} / \lambda_j^{(i)} \in \partial g_j^{(i)}(w_j^{(i)})$  for  $j \in J_i \neq \emptyset$ .

*Proof.* Let  $f \in \mathcal{L}_H$  be any convex function such that  $f(x) = f_1(x) + \dots + f_N(x)$  where  $f_i$  is homogeneous of  $p_i$ -degree, and let  $f$  satisfy condition (4.1). Let  $y^*$  be any element in  $\text{dom}(f^*) = \mathcal{RD}(f)$ . This implies that there exists some point  $x^* \in \text{dom}(f)$  such that  $y^* \in \partial f(x^*)$ . Then, for any  $x \in \text{dom}(f)$ , we have  $f(x) \geq f(x^*) + (y^*)^T(x - x^*)$  which can be written as  $(y^*)^T x - f(x) \leq (y^*)^T x^* - f(x^*)$  for all  $x \in \text{dom}(f)$ . This together with Lemma 4.1 implies that

$$(4.5) \quad f^*(y^*) = (y^*)^T x^* - f(x^*) = \sum_{i=1}^N p_i f_i(x^*) - f(x^*) = \sum_{i=1}^N (p_i - 1) f_i(x^*).$$

Setting  $f = g_j^{(i)}$  and  $y^* = u_j^{(i)} / \lambda_j^{(i)}$ , where  $g_j^{(i)}$  is given by (4.3), it follows from (4.5) that

$$\left( g_j^{(i)} \right)^* (u_j^{(i)} / \lambda_j^{(i)}) = \sum_{k=1}^{m^{(ij)}} (p_k^{(ij)} - 1) h_k^{(ij)}(w_j^{(i)})$$

where  $w_j^{(i)}$  can be any point such that  $u_j^{(i)} / \lambda_j^{(i)} \in \partial g_j^{(i)}(w_j^{(i)})$ . Substituting the above into Theorem 3.3, we have the desired result.  $\square$

We now consider the case in which all the function  $g_j^{(i)}(j = 1, \dots, \ell^{(i)})$  are homogeneous. This is a special case of (4.3) with  $m^{(ij)} = 1$  (for all  $j = 1, \dots, \ell^{(i)}, i = 1, \dots, m$ ). We have the following result.

**COROLLARY 4.3.** *Let  $K_i$  be given by (3.4) where each  $g_j^{(i)}(j = 1, \dots, \ell^{(i)})$  is convex and homogeneous of  $p_j^{(i)}$ -degree, and  $g_j^{(i)}$  satisfy Assumption 4.1. Then the robust programming problem (3.2) is equivalent to (4.4), but  $\Upsilon^{(i)}$  is given as follows*

$$\Upsilon^{(i)} = \begin{cases} \sum_{j \in J_i} (p_j^{(i)} - 1) \lambda_j^{(i)} g_j^{(i)}(w_j^{(i)}), & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise} \end{cases},$$

where  $w_j^{(i)}$  satisfies that  $u_j^{(i)}/\lambda_j^{(i)} \in \partial g_j^{(i)}(w_j^{(i)})$  for  $j \in J_i \neq \emptyset$ .

It is worth mentioning that  $\Upsilon^{(i)}$  can be written as

$$\Upsilon^{(i)} = \begin{cases} \sum_{j \in J_i} \left(1 - 1/p_j^{(i)}\right) (u_j^{(i)})^T w_j^{(i)}, & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases}$$

This follows from (ii) of Lemma 4.1. Actually, for any function  $f$  satisfying Assumption 4.1, (4.5) can also be written as  $f^*(y^*) = (y^*)^T x^* - f(x^*) = (1 - 1/p)(y^*)^T x^*$ . Therefore,

$$\left(g_j^{(i)}\right)^* (u_j^{(i)}/\lambda_j^{(i)}) = (1 - 1/p_j^{(i)})(u_j^{(i)})^T w_j^{(i)}/\lambda_j^{(i)}$$

for some  $w_j^{(i)}$  such that  $u_j^{(i)}/\lambda_j^{(i)} \in \partial g_j^{(i)}(w_j^{(i)})$ .

*Remark 4.1.* (i) Notice that in Corollary 4.3 we do not require Slater's condition, since it was shown in [32] that for homogeneous convex optimization, Lagrangian duality results hold without Slater's condition. (ii) It should be mentioned that Slater's condition in Theorem 4.2 is not essential, and can be removed in many situations, or enforced by slightly changing the constants  $\Delta_j^{(i)}$  in (3.4). Any function  $g$  in the linear space  $\mathcal{L}_H$  is the sum of some homogeneous functions whose value is zero at the origin. Thus  $0 \in K_i$  implies that  $0 = g_j^{(i)}(0) \leq \Delta_j^{(i)}$  for  $j = 1, \dots, \ell^{(i)}$ , i.e., all constants  $\Delta_j^{(i)}$  must be nonnegative in (3.4) when  $g_j^{(i)} \in \mathcal{L}_H$ . If all  $\Delta_j^{(i)}$  are positive, Slater's condition holds trivially (this is the situation in most practical applications; for example, when  $g_j^{(i)}$  is a norm,  $\Delta_j^{(i)}$  is positive since otherwise the uncertainty set contains at most one point). If not all  $\Delta_j^{(i)}$  are positive, replacing  $\Delta_j^{(i)}$  in (3.4) by  $\hat{\Delta}_j^{(i)}$  where  $\hat{\Delta}_j^{(i)} = \Delta_j^{(i)}$  if  $\Delta_j^{(i)} > 0$ , and  $\hat{\Delta}_j^{(i)} = \varepsilon$  otherwise, for some small  $\varepsilon > 0$ , allows to satisfy Slater's condition.

In the next section, we show that in homogeneous cases the above results can be further improved without making Assumption 4.1.

**5. Homogeneous cases.** We now show that for homogeneous of 1-degree functions, Assumption 4.1 holds trivially, and for a degree  $p \neq 1$ , a simple transformation will make the resulting functions satisfy Assumption 4.1. We also further simplify the reformulation. We first prove some basic properties of homogeneous functions. Part (i) of the following lemma in fact follows from [30], but for completeness we provide a simple proof. It appears that the result of part (ii) of the following lemma should be valid for non-differentiable functions as well, but for simplicity of the proof we state it for twice differentiable functions.

**LEMMA 5.1.** *Let  $f : \text{dom}(f) \subseteq R^n \rightarrow R$  be convex and homogeneous of  $p$ -degree.*

*(i) If the degree  $p > 1$ , then  $f(x) \geq 0$  over its domain, and if  $p < 1$ , then  $f(x) \leq 0$  over its domain.*

*(ii) Let  $f$  be twice differentiable over its domain. Then for  $p > 1$ , the function  $(f(x))^{1/p}$  is convex and homogeneous of 1-degree; For  $p < 1$ , the function  $-(-f(x))^{1/p}$  is convex and homogeneous of 1-degree.*

*Proof.* Let  $x$  be any point in  $\text{dom}(f)$ . By homogeneity and convexity of  $f$ , we have

$$(1/2)^p f(x) = f(x/2) \leq f(x)/2 + f(0)/2 = f(x)/2.$$

Thus,  $[(1/2)^p - 1/2] f(x) \leq 0$ , and hence the result (i) follows.



We now prove the result of part (ii). Consider the case of  $p > 1$ . By (i),  $p > 1$  implies that  $f(x) \geq 0$  over its domain. Let  $\varepsilon > 0$  be any given positive number. Denote  $g_\varepsilon(x) := (f(x) + \varepsilon)^{1/p}$ . Notice that  $\text{dom}(g_\varepsilon) = \text{dom}(f)$ , and  $g_\varepsilon$  is twice differentiable. We prove first that  $g_\varepsilon$  is a convex function for any given  $\varepsilon > 0$ . It suffices to show that  $\nabla^2 g_\varepsilon(x) \succeq 0$  (positive semi-definite). Since

$$\nabla^2 g_\varepsilon(x) = \frac{1}{p}(f(x) + \varepsilon)^{\frac{1}{p}-2} \left[ \left( \frac{1}{p} - 1 \right) \nabla f(x) \nabla f(x)^T + (f(x) + \varepsilon) \nabla^2 f(x) \right],$$

it is sufficient to prove that

$$\left( \frac{1}{p} - 1 \right) \nabla f(x) \nabla f(x)^T + (f(x) + \varepsilon) \nabla^2 f(x) \succeq 0.$$

By Schur complementarity property, this is equivalent to showing that

$$\begin{bmatrix} \frac{p}{p-1}(f(x) + \varepsilon) & \nabla f(x)^T \\ \nabla f(x) & \nabla^2 f(x) \end{bmatrix} \succeq 0.$$

Thus, we need to show for all  $(t, u) \in R^{n+1}$  that

$$\begin{aligned} \varphi(t, u) &= (t, u^T) \begin{bmatrix} \frac{p}{p-1}(f(x) + \varepsilon) & \nabla f(x)^T \\ \nabla f(x) & \nabla^2 f(x) \end{bmatrix} \begin{pmatrix} t \\ u \end{pmatrix} \\ &= \frac{p}{p-1} t^2 (f(x) + \varepsilon) + 2t \nabla f(x)^T u + u^T \nabla^2 f(x) u \geq 0. \end{aligned}$$

Case 1:  $t = 0$ . By convexity of  $f$ ,  $u^T \nabla^2 f(x) u \geq 0$  for any  $u \in R^n$ , thus we have  $\varphi(t, u) \geq 0$ .

Case 2:  $t \neq 0$ . In this case, it suffices to show that for any  $u \in R^n$

$$\varphi(1, u) = \frac{p}{p-1} (f(x) + \varepsilon) + 2 \nabla f(x)^T u + u^T \nabla^2 f(x) u \geq 0.$$

Since  $\nabla^2 f(x) \succeq 0$ , the function  $\varphi(1, u)$  is convex with respect to  $u$ , and its minimum is attained if there exists some  $u^*$  such that

$$(5.1) \quad \nabla f(x) = -\nabla^2 f(x) u^*,$$

and the minimum value is

$$\varphi(1, u^*) = \frac{p}{p-1} (f(x) + \varepsilon) + \nabla f(x)^T u^*.$$

By Euler's formula, we have  $x^T \nabla f(x) = p f(x)$ . Differentiating both sides of this equation, we have  $(p-1) \nabla f(x) = \nabla^2 f(x) x$ , which shows that the vector  $u^* = -\frac{1}{p-1} x$  satisfies equation (5.1), thus the minimum

$$\varphi(1, u^*) = \frac{p}{p-1} (f(x) + \varepsilon) - \frac{1}{p-1} \nabla f(x)^T x = \frac{p}{p-1} \varepsilon > 0.$$

The last equation follows from Euler's formula again. Therefore  $\varphi(t, u) \geq 0$  for any  $(t, u) \in R^{n+1}$ . Convexity of  $g_\varepsilon(x)$  follows. Since  $\varepsilon > 0$  is arbitrary and  $(f(x))^{1/p} = \lim_{\varepsilon \rightarrow 0} g_\varepsilon(x)$ , we conclude that  $(f(x))^{1/p}$  is convex.

The case of  $p < 1$  is considered analogously.  $\square$

According to our definition of a homogeneous function, its domain includes the origin. The next lemma shows that Assumption 4.1 is satisfied for any homogeneous of 1-degree convex function, and its subdifferential at the origin defines the domain of the conjugate function.

LEMMA 5.2. *Let  $h : \text{dom}(h) \subseteq \mathbb{R}^N \rightarrow \overline{\mathbb{R}}$  be a closed proper convex function and be homogeneous of 1-degree. Then*

$$\mathcal{RD}(h) = \bigcup_{x \in \text{dom}(h)} \partial h(x) = \partial h(0).$$

Moreover,  $\text{dom}(h^*) = \mathcal{RD}(h) = \partial h(0)$ .

*Proof.* . Let  $z$  be any subgradient of  $h$  at  $x$ , then for any given  $y$  and any positive number  $\lambda$  we have  $h(\lambda y) \geq h(x) + z^T(\lambda y - x)$ . Since  $\lambda$  is positive, dividing both sides of the inequality by  $\lambda$ , and using homogeneity of  $h$ , we have

$$h(y) \geq \frac{h(x) - z^T x}{\lambda} + z^T y.$$

Let  $\lambda \rightarrow \infty$ . We have  $h(y) \geq z^T y$  which holds for any  $y$ . Consider the set:

$$S := \{z : z^T y \leq h(y) \text{ for any } y \in \text{dom}(h)\}.$$

From the above proof, we have seen that  $\partial h(x) \subseteq S$  for any  $x$ , i.e.,  $\mathcal{RD}(h) \subseteq S$ . In particular, we have  $\partial h(0) \subseteq S$ . Conversely, since  $h(0) = 0$ , we see that any  $z \in S$  is a subgradient of  $h$  at  $x = 0$ . Thus, we have  $S \subseteq \partial h(0)$ . We conclude that  $\mathcal{RD}(h) = S = \partial h(0)$ . The first part of the lemma has been proved.

We now prove the second part of the lemma. For any  $y^* \in \mathcal{RD}(h)$ , there exists an  $x^*$  such that  $y^* \in \partial f(x^*)$ , and by definition of sub-gradient, we have that  $(y^*)^T x - h(x) \leq (y^*)^T x^* - h(x^*)$  for any  $x \in \text{dom}(h)$ , which implies that  $h^*(y^*) < \infty$ , i.e.,  $y^* \in \text{dom}(h^*)$ . Thus, the inclusion  $\mathcal{RD}(h) \subseteq \text{dom}(h^*)$  holds trivially (we mentioned this observation at the beginning of Section 4).

Now we show that converse inclusion is also valid. Suppose that  $y^* \in \text{dom}(h^*)$ . We show that  $y^* \in S$ . Notice that for homogeneous of 1-degree function  $h$ ,  $\text{dom}(h)$  is a cone. Thus, for any given positive number  $\lambda$ , we have

$$\lambda h^*(y^*) = \sup_{x \in \text{dom}(h)} (y^*)^T(\lambda x) - \lambda h(x) = \sup_{x \in \text{dom}(h)} (y^*)^T(\lambda x) - h(\lambda x) = h^*(y^*)$$

since  $\lambda > 0$  can be any positive number, we have  $h^*(y^*) = 0$ , which in turn implies that  $(y^*)^T x - h(x) \leq h^*(y^*) = 0$  for any  $x \in \text{dom}(h)$ , and therefore  $y^* \in S$ . The desired result follows.  $\square$

We can now simplify the robust counterpart for the homogeneous 1-degree case.

THEOREM 5.3. *Let  $K_i$  be defined by (3.4) where all the functions  $g_j^{(i)}$ ,  $i = 1, \dots, \ell^{(i)}$ , are closed proper convex functions and are homogeneous of 1-degree. Then the robust counterpart (3.2) is equivalent to*

$$\begin{aligned} & \min c^T x \\ & \text{s.t. } \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} \leq b_i, \quad i = 1, \dots, m, \\ (5.2) \quad & \chi_i = \begin{cases} \sum_{j \in J_i} u_j^{(i)}, & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases} \quad i = 1, \dots, m, \end{aligned}$$

$$\begin{aligned}\lambda_j^{(i)} &\geq 0, \quad j = 1, \dots, \ell^{(i)}, \quad i = 1, \dots, m, \\ F(x) &\leq 0,\end{aligned}$$

where  $\chi_i$  and  $J_i$  are the same as in Theorem 4.1, and  $u_j^{(i)}/\lambda_j^{(i)} \in \partial g_j^{(i)}(0)$  for  $j \in J_i \neq \emptyset$ ,  $i = 1, \dots, m$ .

*Proof.* Under the conditions of the theorem, Lemma 5.2 claims that the Assumption 4.1 holds, and moreover  $\mathcal{RD}(g_j^{(i)}) = \partial g_j^{(i)}(0)$  for all  $i = 1, \dots, \ell^{(i)}$ . From the proof of Theorem 4.2, when  $J_i \neq \emptyset$ , we can set  $w_j^{(i)} = 0$ , and hence  $\Upsilon^{(i)} = \left(g_j^{(i)}\right)^* (u_j^{(i)}/\lambda_j^{(i)}) = 0$ . Thus, in this case,  $\Upsilon^{(i)} \equiv 0$  no matter what  $J_i$  is. Therefore, the robust counterpart (3.2) eventually reduces to (5.2). As mentioned in Remark 4.1, we do not need Slater's condition for homogeneous cases.  $\square$

When  $g_j^{(i)}$  is homogeneous of  $p_j^{(i)}$ -degree where  $p_j^{(i)} \neq 1$  and twice differentiable, by (ii) of Lemma 5.1, we may transform it into a homogeneous of 1-degree function. Then, we can use Theorem 5.3. When  $p_j^{(i)} < 1$ , by Lemma 5.1, the value of  $g_j^{(i)}$  is non-positive, thus the constraint  $g_j^{(i)} \leq \Delta_j^{(i)}$  becomes redundant (since  $\Delta_j^{(i)} \geq 0$ ) and thus can be removed from the list of constraints defining  $K_i$ . Therefore, without loss of generality, we assume that all  $p_j^{(i)} \geq 1$ . We now have the following result.

**THEOREM 5.4.** *Let  $K_i$  be defined by (3.4) where the functions  $g_j^{(i)}$ ,  $j = 1, \dots, \ell^{(i)}$ , are twice differentiable, convex and homogeneous of  $p_j^{(i)}$ -degree ( $p_j^{(i)} \geq 1$ ), respectively. Then, the robust programming problem (3.2) is equivalent to*

$$\begin{aligned}\min & c^T x \\ \text{s.t.} & \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \widetilde{\Delta_j^{(i)}} \leq b_i, \quad i = 1, \dots, m, \\ & \chi_i = \begin{cases} \sum_{j \in J_i} u_j^{(i)} & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases} \quad i = 1, \dots, m, \\ & \lambda_j^{(i)} \geq 0, j = 1, \dots, \ell^{(i)}, \quad i = 1, \dots, m, \\ & F(x) \leq 0,\end{aligned}$$

where  $\chi_i$  and  $J_i$  are the same as in Theorem 4.1,  $u_j^{(i)} \in \lambda_j^{(i)} \partial \mathcal{G}_j^{(i)}(0)$  for  $j \in J_i \neq \emptyset$ ,  $i = 1, \dots, m$  and

$$(5.3) \quad \mathcal{G}_j^{(i)} = \begin{cases} (g_j^{(i)})^{1/p_j^{(i)}}, & p_j^{(i)} > 1, \\ g_j^{(i)}, & p_j^{(i)} = 1 \end{cases}, \quad \widetilde{\Delta_j^{(i)}} = \begin{cases} (\Delta_j^{(i)})^{1/p_j^{(i)}}, & p_j^{(i)} > 1, \\ \Delta_j^{(i)}, & p_j^{(i)} = 1. \end{cases}$$

*Proof.* We note that for  $p_j^{(i)} > 1$ , since  $g_j^{(i)}$  and  $\Delta_j^{(i)}$  are nonnegative by Lemma 5.1, the constraint  $g_j^{(i)} \leq \Delta_j^{(i)}$  in (3.4) is equivalent to  $(g_j^{(i)})^{1/p_j^{(i)}} \leq (\Delta_j^{(i)})^{1/p_j^{(i)}}$ . Define  $\mathcal{G}_j^{(i)}$  and  $\widetilde{\Delta_j^{(i)}}$  as in (5.3). Then this result is an immediate consequence of Theorem 5.3 and Lemma 5.1.  $\square$

From Theorems 5.3 and 5.4, the structure of robust counterparts of uncertain optimization problems mainly depends on the subdifferentials of  $g_j^{(i)}$  or  $\mathcal{G}_j^{(i)}$  at the origin when functions  $g_j^{(i)}$  are homogeneous.

Notice that any norm is convex and homogeneous of 1-degree and can be defined on the whole space (but the converse is not true, for example consider  $f(t) : R \rightarrow R$

given by  $f(t) = t$  if  $t \geq 0$  and  $f(t) = 2|t|$  if  $t < 0$ . Clearly,  $f$  is convex and homogeneous of 1-degree, but it is not a norm, because  $f(-1) \neq f(1)$ . Theorem 5.3 can be immediately applied to the case of an uncertainty set defined by a finite system of norm inequalities. For this case, however, in addition to the above formulation of the robust counterpart via subgradients at the origin, we can further simplify it using dual norms and eliminating all variables  $\lambda_j^{(i)}$ . For any norm  $\|\cdot\|$ , we denote its dual norm by  $\|\cdot\|_*$ , i.e.,  $\|u\|_* = \sup_{\|x\| \leq 1} u^T x$ . When  $g_j^{(i)}$  is a norm, we denote it by  $\|\cdot\|^{(ij)}$ , and its dual norm by  $\|\cdot\|_*^{(ij)}$ .

**COROLLARY 5.5.** *Let  $K_i$  be defined by (3.4) where all  $g_j^{(i)}$  ( $j = 1, \dots, \ell^{(i)}, i = 1, \dots, m$ ) are norms, denoted respectively by  $\|\cdot\|^{(ij)}$  ( $j = 1, \dots, \ell^{(i)}, i = 1, \dots, m$ ), then the robust counterpart (3.2) is equivalent to*

$$\begin{aligned} & \min c^T x \\ & \text{s.t. } \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \sum_{j=1}^{\ell^{(i)}} \Delta_j^{(i)} \left\|u_j^{(i)}\right\|_*^{(ij)} \leq b_i, \quad i = 1, \dots, m, \\ & \chi_i = \sum_{j=1}^{\ell^{(i)}} u_j^{(i)}, \quad i = 1, \dots, m, \\ & F(x) \leq 0. \end{aligned}$$

where  $\chi_i = W^{(i)}(x) \otimes V^{(i)}(x)$ .

*Proof.* Notice that  $u \in \partial\|0\|$  if and only if  $u^T x \leq \|x\|$  for any  $x$  which can be written as  $u^T(x/\|x\|) \leq 1$ , i.e.,  $\|u\|_* \leq 1$ . Therefore, for  $j \in J_i \neq \emptyset$ ,  $u_j^{(i)}/\lambda_j^{(i)} \in \partial g_j^{(i)}(0)$  is equivalent to  $\left\|\frac{u_j^{(i)}}{\lambda_j^{(i)}}\right\|_*^{(ij)} \leq 1$ , or just  $\|u_j^{(i)}\|_*^{(ij)} \leq \lambda_j^{(i)}$ . Therefore, the constraints of (5.2) can be further written as

$$\begin{aligned} & \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} \leq b_i, \quad i = 1, \dots, m, \\ & \chi_i = \begin{cases} \sum_{j \in J_i} u_j^{(i)}, & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise} \end{cases}, \quad i = 1, \dots, m, \\ & \lambda_j^{(i)} \geq 0, \quad j = 1, \dots, \ell^{(i)}; \quad i = 1, \dots, m, \\ & \|u_j^{(i)}\|_*^{(ij)} \leq \lambda_j^{(i)}, \quad \forall j \in J_i \neq \emptyset, \quad i = 1, \dots, m, \\ & F(x) \leq 0, \end{aligned}$$

It is evident that the above system is equivalent to

$$\begin{aligned} & \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \sum_{j=1}^{\ell^{(i)}} \lambda_j^{(i)} \Delta_j^{(i)} \leq b_i, \quad i = 1, \dots, m, \\ & \|u_j^{(i)}\|_*^{(ij)} \leq \lambda_j^{(i)}, \quad j = 1, \dots, \ell^{(i)}; \quad i = 1, \dots, m, \\ & \chi_i = \sum_{j=1}^{\ell^{(i)}} u_j^{(i)}, \quad i = 1, \dots, m, \\ & \lambda_j^{(i)} \geq 0, \quad j = 1, \dots, \ell^{(i)}; \quad i = 1, \dots, m, \end{aligned}$$

$$F(x) \leq 0,$$

Eliminating the variables  $\lambda_j^{(i)}$ , the above system becomes

$$\begin{aligned} \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \sum_{j=1}^{\ell^{(i)}} \Delta_j^{(i)} \left\| u_j^{(i)} \right\|_*^{(ij)} &\leq b_i, \quad i = 1, \dots, m, \\ \chi_i &= \sum_{j=1}^{\ell^{(i)}} u_j^{(i)}, \quad i = 1, \dots, m, \\ F(x) &\leq 0. \end{aligned}$$

The desired result is obtained.  $\square$

**6. Special cases.** Complexity of robust counterparts depends both on the structure of the original optimization problems and on the structure of the uncertainty set. The harder the original optimization problem is and/or the more complex the uncertainty set is, the more difficult the robust counterpart is. In this section, we demonstrate how the general results developed above can be simplified by considering special optimization problems and/or special uncertainty sets. We take linear programming problem (LP) as an example of a special optimization problem, and take the widely used uncertainty set (3.5) as an example of a special uncertainty set. Thus we obtain new results for problem (2.1) with uncertainty set defined by (3.5) and for robust LP with general uncertainty sets. For this simplest of the considered cases (robust LP with uncertainty set (3.5)), we show that our results contain a number of related results in the literature, but under less restrictive assumptions, thus generalizing and strengthening these results.

**6.1. Problem (2.1) with uncertainty set  $\mathcal{U}$  defined by (3.5).** Now we consider the uncertainty set (3.5), i.e.,

$$\mathcal{U} =: \left\{ D \left| \exists z \in R^{|N|} : D = D_0 + \sum_{j \in N} \Delta D_j z_j, \quad \|z\| \leq \Omega \right. \right\}.$$

Since this model has been widely used in the literature (see for instance, [5]-[14]), it is interesting to see how our general results can be simplified when reduced to the above uncertainty set. Let  $H$  denote the matrix whose columns are  $\Delta D_j, j = 1, \dots, |N|$ , i.e.,

$$H = [\Delta D_1, \dots, \Delta D_{|N|}].$$

Define the function

$$(6.1) \quad g(u) = \inf \{ \|z\| : Hz = u \}.$$

Then  $g(u)$  is convex and homogeneous of 1-degree (convexity is proven in [36], and homogeneity can be checked directly). Now we show that the uncertainty set (3.5) can be represented equivalently in the form (3.4).

**LEMMA 6.1.** *Consider the uncertainty set  $\mathcal{U}$  given by (3.5). Let  $K = \{u \mid g(u) \leq \Omega\}$ , where  $g$  is given by (6.1). Then we have  $K = \mathcal{U} - \{D_0\}$ .*

*Proof.* Let  $u$  be any point in  $K$ . By the definition of  $g(u)$ , there exists a point  $z^*$  such that  $g(u) = \|z^*\|$  and  $H z^* = u$ . Since  $u \in K$  implies  $g(u) \leq \Omega$ , we have  $\|z^*\| \leq \Omega$ . By the definition of  $\mathcal{U}$ , we see that  $u \in \mathcal{U} - \{D_0\}$ .

Conversely, suppose that  $u \in \mathcal{U} - \{D_0\}$ . Then there exists a point  $D \in \mathcal{U}$  such that  $u = D - D_0$ . By the definition of  $\mathcal{U}$ , there exists a point  $z$  such that  $u = Hz$  and  $\|z\| \leq \Omega$ . By the definition of  $g$ , this implies  $g(u) \leq \Omega$ , and hence  $u \in K$ .  $\square$

If the vectors  $\{\Delta D_j : j = 1, \dots, N\}$  are linearly independent, from  $Hz = u$  we have  $z = (H^T H)^{-1} H^T u$ . Thus, we have

$$\mathcal{U} - \{D_0\} = K = \{u \mid g(u) = \|(H^T H)^{-1} H^T u\| \leq \Omega\}.$$

Since in general  $|N|$  is less than the number of data of the problem, the term  $H^T u$  can be zero even when  $u \neq 0$ . Thus,  $g(u)$  is not a norm in this case, unless  $\{\Delta D_j : j = 1, \dots, N\}$  are linearly independent and  $|N|$  equals to the number of data of the problem, in which case  $H$  is an  $|N| \times |N|$  invertible matrix.

Notice that  $K$  here has only one constraint which corresponds to the case  $\ell^{(i)} = 1$  for all  $i = 1, \dots, m$ , and by Theorem 16.3 in [36] the conjugate function of  $g(u)$  is given by

$$(6.2) \quad g^*(w) = \begin{cases} 0, & \|H^T w\|_* \leq 1, \\ \infty, & \text{otherwise,} \end{cases}$$

where  $\|\cdot\|_*$  denotes the dual norm of  $\|\cdot\|$ .

We now consider our problem (2.1) where data  $M^{(i)}$ 's are subject to uncertainty of the type (3.5), i.e., for each  $i$ , the data  $M^{(i)}$  belong to the set

$$(6.3) \quad \left\{ M^{(i)} \mid \exists z \in R^{|N^{(i)}|} : M^{(i)} = \overline{M}_0^{(i)} + \sum_{j \in N^{(i)}} \Delta M_j^{(i)} z_j, \quad \|z\|^{(i)} \leq \Omega^{(i)} \right\}.$$

This can be equivalently written as

$$(6.4) \quad \mathcal{U}_i = \left\{ \text{vec}(M^{(i)}) \mid \exists z \in R^{|N^{(i)}|} : \text{vec}(M^{(i)}) = \text{vec}(\overline{M}_0^{(i)}) + \sum_{j \in N^{(i)}} \text{vec}(\Delta M_j^{(i)}) z_j, \quad \|z\|^{(i)} \leq \Omega^{(i)} \right\},$$

$i = 1, \dots, m$ , where  $N^{(i)}$  is the corresponding index set (not to be confused with  $N_i$  - the dimension of matrix  $M^{(i)}$ ), and  $\Omega^{(i)}$  is a given number. Note that we add the index  $(i)$  to the norm (i.e.,  $\|\cdot\|^{(i)}$ ), which allows us to use different norms for different constraints. Accordingly, we have the function

$$g^{(i)}(u) = \inf\{\|z\|^{(i)} : H^{(i)} z = u\},$$

where  $H^{(i)} = [\text{vec}(\Delta M_1^{(i)}), \text{vec}(\Delta M_2^{(i)}), \dots, \text{vec}(\Delta M_{|N^{(i)}|}^{(i)})]$  and thus by Lemma 6.1 we have

$$\mathcal{U}_i - \{\text{vec}(\overline{M}^{(i)})\} = K_i = \{u \mid g^{(i)}(u) \leq \Omega^{(i)}\}.$$

Using (6.2), we have

$$(6.5) \quad (g^{(i)})^*(w) = \begin{cases} 0, & \|(H^{(i)})^T w\|_*^{(i)} \leq 1 \\ \infty, & \text{otherwise.} \end{cases}$$

Now we have all necessary ingredients to develop our result. We first note that in this case,  $\ell^{(i)} = 1$  for all  $i = 1, \dots, m$  since the uncertainty set  $\mathcal{U}_i$  has only one constraint

$g^{(i)}(u) \leq \Omega^{(i)}$ . So,  $\lambda^{(i)}$  is reduced to a scalar. Therefore, the constraints of the robust counterpart (3.7) that correspond to index  $i$  reduce to

$$(6.6) \quad \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \lambda^{(i)} \Omega^{(i)} + \Upsilon^{(i)} \leq b_i,$$

$$(6.7) \quad \chi_i = \begin{cases} u^{(i)}, & \lambda^{(i)} > 0, \\ 0, & \lambda^{(i)} = 0, \end{cases}$$

where

$$(6.8) \quad \Upsilon^{(i)} = \begin{cases} \lambda^{(i)} (g^{(i)})^* (u^{(i)} / \lambda^{(i)}), & \lambda^{(i)} > 0, \\ 0, & \lambda^{(i)} = 0. \end{cases}$$

When  $\lambda^{(i)} > 0$ , the system (6.6)-(6.8) becomes

$$\left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \lambda^{(i)} \Omega^{(i)} + \lambda^{(i)} (g^{(i)})^* (u^{(i)} / \lambda^{(i)}) \leq b_i,$$

$$\chi_i = u^{(i)}.$$

Eliminating  $u^{(i)}$  and using (6.5), the above system is equivalent to

$$\left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \lambda^{(i)} \Omega^{(i)} \leq b_i,$$

$$\|(H^{(i)})^T (\chi_i)\|_*^{(i)} \leq \lambda^{(i)}.$$

This can be written as

$$(6.9) \quad \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \Omega^{(i)} \|(H^{(i)})^T \chi_i\|_*^{(i)} \leq b_i.$$

When  $\lambda^{(i)} = 0$ , the system (6.6)-(6.8) is written as

$$\left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) \leq b_i,$$

$$\chi_i = 0.$$

Clearly, this system can be written as (6.9), too. Hence, by Theorem 3.3, we have the following result.

**THEOREM 6.2.** *Under the uncertainty set (6.3) (or equally, (6.4)), the robust counterpart (3.2) is equivalent to*

$$\begin{aligned} & \min c^T x \\ & s.t. \quad \left(W^{(i)}(x)\right)^T \overline{M}^{(i)} V^{(i)}(x) + \Omega^{(i)} \|(H^{(i)})^T \chi_i\|_*^{(i)} \leq b_i, \quad i = 1, \dots, m, \\ & \quad F(x) \leq 0, \end{aligned}$$

where  $\chi_i = W^{(i)}(x) \otimes V^{(i)}(x)$  and  $H^{(i)} = \left[ \text{vec}(\Delta M_1^{(i)}), \text{vec}(\Delta M_2^{(i)}), \dots, \text{vec}(\Delta M_{N^{(i)}}^{(i)}) \right]$

**6.2. Linear programming with general uncertainty sets.** Consider the LP problem discussed in Section 2:  $\min\{c^T x : Ax \leq b, x \geq 0\}$ , where  $A \in R^{m \times n}$ ,  $b \in R^m$  and  $c \in R^n$ . As discussed in Section 2, without loss of generality, we assume that only the coefficients of  $A$  are subject to uncertainty.

There are two widely used ways to characterize the uncertain data of LP problems. One is the “row-wise” uncertainty model (a separate uncertainty set is specified for each row of  $A$ ), and the other is what we may call the “global” uncertainty model (one uncertainty set for the whole matrix  $A$  is specified). We first consider the situation of “global” uncertainty.

Suppose that  $A$  is allowed to vary in such a way that its deviations from a given nominal  $\bar{A}$  fall in a bounded convex set  $K$  of  $R^{mn}$  that contains the origin (zero). That is, the uncertainty set is defined as

$$(6.10) \quad \mathcal{U} = \{\tilde{A} | \text{vec}(\tilde{A}) - \text{vec}(\bar{A}) \in K\},$$

where  $K$  is defined by convex inequalities:

$$(6.11) \quad K = \{u | g_j(u) \leq \Delta_j, j = 1, \dots, \ell\}.$$

Here  $\Delta_j$ 's are constants, and all  $g_j$  are closed proper convex functions. Then the robust counterpart of the LP problem with uncertainty set  $\mathcal{U}$  is

$$(6.12) \quad \min\{c^T x : \tilde{A}x \leq b, x \geq 0, \forall \tilde{A} \in \mathcal{U}\}.$$

First of all, from Section 2 we know that for LP we can drop indexes  $i$  for  $g_j^{(i)}$  and  $\Delta_j^{(i)}$  in the previous discussion, since in the reformulation of LP as a special case of (2.1) and (2.2), the data matrix for each constraint (2.2) is the same, i.e.,  $M^{(i)} = \begin{bmatrix} A \\ 0 \end{bmatrix}_{n \times n}$  for all  $i$  (see Section 2). Second, we note that for LP, the vector  $\chi_i = W^{(i)} \otimes V^{(i)} = e_i \otimes x$  is linear in  $x$ . Therefore, the results in previous sections can be further simplified for LP. For example, Theorems 3.1, 3.3, 5.4, and Corollary 5.5 can be stated as follows (Theorems 6.3 through 6.5 and Corollary 6.6, respectively).

**THEOREM 6.3.** *The robust LP problem (6.12) is equivalent to the convex programming problem*

$$\begin{aligned} & \min c^T x \\ & \text{s.t. } \bar{a}_i^T x + \delta^*(\chi_i | \text{cl}(\text{co}(K))) \leq b_i, \quad i = 1, \dots, m, \\ & \quad x \geq 0, \end{aligned}$$

where  $\text{cl}(\text{co}(K))$  is the closed convex hull of the set  $K$ , and  $\chi_i = e_i \otimes x$ .

Since  $\delta^*(\cdot | \text{cl}(\text{co}(K)))$  is a closed convex function, the robust counterpart of any LP problem with the uncertainty set denoted by (6.10) and (6.11) is a convex programming problem.

**THEOREM 6.4.** *Let  $K$  be given by (6.11) where  $g_j (j = 1, \dots, \ell)$  are arbitrary closed proper convex functions. Suppose that Slater's condition holds, i.e., there exists a point  $u_0$  such that  $g_j(u_0) < \Delta_j$  for all  $j = 1, \dots, \ell$ . Then the robust LP problem (6.12) is equivalent to*

$$\begin{aligned} & \min c^T x \\ & \text{s.t. } \bar{a}_i^T x + \sum_{j=1}^{\ell} \lambda_j^{(i)} \Delta_j + \left( \sum_{j=1}^{\ell} \lambda_j^{(i)} g_j \right)^* (\chi_i) \leq b_i, \quad i = 1, \dots, m, \end{aligned}$$



$$\begin{aligned}\lambda_j^{(i)} &\geq 0, \quad j = 1, \dots, \ell; \quad i = 1, \dots, m, \\ x &\geq 0,\end{aligned}$$

or equivalently

$$\begin{aligned}(6.13) \quad & \min c^T x \\ & \text{s.t. } \bar{a}_i^T x + \sum_{j=1}^{\ell} \lambda_j^{(i)} \Delta_j + \Upsilon^{(i)} \leq b_i, \quad i = 1, \dots, m, \\ & \chi_i = \begin{cases} \sum_{j \in J_i} u_j^{(i)}, & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise} \end{cases}, \quad i = 1, \dots, m, \\ & \lambda_j^{(i)} \geq 0, \quad j = 1, \dots, \ell; \quad i = 1, \dots, m, \\ & x \geq 0,\end{aligned}$$

where  $\chi_i = e_i \otimes x$ , and  $J_i = \{j : \lambda_j^{(i)} > 0, j = 1, \dots, \ell\}$ , and

$$\Upsilon^{(i)} = \begin{cases} \sum_{j \in J_i} \lambda_j^{(i)} g_j^*(u_j^{(i)} / \lambda_j^{(i)}), & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases}$$

*Remark 6.1.* (i) For LP, the constraint “ $\chi_i = \sum_{j \in J_i} u_j^{(i)}$ ” is a linear constraint. (ii) It is well known that for any convex function  $f$ , the function  $\hat{f}(x, t) = tf(x/t)$ , where  $t > 0$ , is also convex in  $(x, t)$ , and is positive homogeneous of 1-degree, that is,  $\hat{f}(\alpha x, \alpha t) = \alpha \hat{f}(x, t)$ , for any  $\alpha > 0$ . Problem (6.13) shows that all functions involved are homogeneous of 1-degree with respect to the variables  $(x, \lambda^{(i)}, u^{(i)})$ . Thus, the robust LP problem (6.12) is not only a convex programming problem, but also a homogeneous programming problem, i.e., an optimization problem where all functions involved are homogeneous.

**THEOREM 6.5.** *Let  $K$  be defined by (6.11) where the functions  $g_j, j = 1, \dots, \ell$ , are twice differentiable, convex and homogeneous of  $p_j$ -degree ( $p_j \geq 1$ ), respectively. Then, the robust LP problem (6.12) is equivalent to*

$$\begin{aligned}& \min c^T x \\ & \text{s.t. } \bar{a}_i^T x + \sum_{j=1}^{\ell} \lambda_j^{(i)} \widetilde{\Delta}_j \leq b_i, \quad i = 1, \dots, m, \\ & \chi_i = \begin{cases} \sum_{j \in J_i} u_j^{(i)}, & \text{if } J_i \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases}, \quad i = 1, \dots, m, \\ & \lambda_j^{(i)} \geq 0, \quad j = 1, \dots, \ell; \quad i = 1, \dots, m, \\ & x \geq 0,\end{aligned}$$

where  $\chi_i$  and  $J_i$  are the same as in Theorem 6.4,  $u_j^{(i)} \in \lambda_j^{(i)} \partial \mathcal{G}_j(0)$  for  $j \in J_i \neq \emptyset, i = 1, \dots, m$  and

$$\mathcal{G}_j = \begin{cases} (g_j)^{1/p_j}, & p_j > 1, \\ g_j, & p_j = 1 \end{cases}, \quad \widetilde{\Delta}_j = \begin{cases} (\Delta_j)^{1/p_j}, & p_j > 1, \\ \Delta_j, & p_j = 1. \end{cases}$$

COROLLARY 6.6. *Let  $K$  be defined by (6.11) where all  $g_j$  ( $j = 1, \dots, \ell$ ) are norms, denoted respectively by  $\|\cdot\|^{(j)}$ ,  $j = 1, \dots, \ell$ , then the robust counterpart (6.12) is equivalent to*

$$\begin{aligned} \min & c^T x \\ \text{s.t.} & \bar{a}_i^T x + \sum_{j=1}^{\ell} \Delta_j \|u_j^{(i)}\|_* \leq b_i, \quad i = 1, \dots, m, \\ & e_i \otimes x = \sum_{j=1}^{\ell} u_j^{(i)}, \quad i = 1, \dots, m, \\ & x \geq 0. \end{aligned}$$

Now we briefly discuss the situation of “row-wise” uncertainty sets. In this case, in order to apply our general results, we reformulate LP in the form (2.1) in a different way than in Section 2. Consider functions  $f_i(x)$  of the form (2.2), where  $W^{(i)}(x) = e_i \in R^n$ ,  $V^{(i)}(x) = x \in R^n$  (same as in Section 2). Throughout the rest of the paper, we denote by  $A_i$  ( $i = 1, \dots, m$ ) the  $i$ th row of  $A$ . Thus,  $A_i$  is an  $n$ -dimensional row vector. The  $n \times n$  matrix  $M^{(i)}$  is the matrix having  $A_i$  as its  $i$ th row and 0 elsewhere, i.e.

$$(6.14) \quad M^{(i)} = \begin{bmatrix} 0 \\ A_i \\ 0 \end{bmatrix}_{n \times n}, \quad i = 1, \dots, m.$$

Then the  $i$ th constraint of  $Ax \leq b$  can be written as

$$(6.15) \quad f_i = (W^{(i)})^T M^{(i)} V^{(i)} \leq b_i$$

for  $i = 1, \dots, m$ . Then applying the results of Sections 3,4,5 to the optimization problem (2.1) with the above inequality constraints and  $F(x) = -x \leq 0$ , we can obtain a formulation for robust LP with “row-wise” uncertainty sets. We omit these results.

The formulation for other special cases such as LCP and QP can be derived similarly; we leave these derivations to interested readers.

**6.3. Linear programming with uncertainty set of type (3.5).** In this section, we consider the LP problem  $\min\{c^T x : Ax \leq b, x \geq 0\}$  under uncertainty of type (3.5). We will show that our results in this section include a number of recent results on robust LP in the literature as special cases. From Theorem 6.2 and 6.4, we have the following result.

THEOREM 6.7. (i) *Under the “row-wise” uncertainty set*

$$(6.16) \quad \mathcal{U}_i = \left\{ A_i \left| \exists u \in R^{N^{(i)}} : A_i = \bar{A}_i + \sum_{j \in N^{(i)}} \Delta A_j^{(i)} u_j, \|u\|^{(i)} \leq \Omega^{(i)} \right. \right\},$$

*the robust counterpart of LP is equivalent to*

$$(6.17) \quad \begin{aligned} \min & c^T x \\ \text{s.t.} & \bar{a}_i^T x + \Omega^{(i)} \left\| \left( H^{(i)} \right)^T x \right\|_*^{(i)} \leq b_i, \quad i = 1, \dots, m, \\ & x \geq 0. \end{aligned}$$

where the matrix  $H^{(i)} = \left[ \left( \Delta A_1^{(i)} \right)^T, \left( \Delta A_2^{(i)} \right)^T, \dots, \left( \Delta A_{|N^{(i)}|}^{(i)} \right)^T \right]$ .

(ii) Under the “global” uncertainty set

$$(6.18) \quad \mathcal{U} = \left\{ A \left| \exists u \in R^{|N|} : A = \bar{A} + \sum_{j \in N} \Delta A_j u_j, \|u\| \leq \Omega \right. \right\},$$

where  $A$  is  $m \times n$  matrix, the robust counterpart of LP is equivalent to

$$(6.19) \quad \begin{aligned} & \min c^T x \\ & \text{s.t. } \bar{a}_i^T x + \Omega \left\| \tilde{H}^T \tilde{\chi}_i \right\|_* \leq b_i, \quad i = 1, \dots, m, \\ & x \geq 0. \end{aligned}$$

where the matrix  $\tilde{H} = [\text{vec}(\Delta A_1), \text{vec}(\Delta A_2), \dots, \text{vec}(\Delta A_{|N|})]$  and  $\tilde{\chi}_i = e_i^{(m)} \otimes x$  where  $e_i^{(m)}$  denotes the  $i$ th column of the  $m \times m$  identity matrix. Equivalently, the inequality (6.19) can be written as

$$\bar{a}_i^T x + \Omega \left\| \left( \tilde{H}^{(i)} \right)^T x \right\|_* \leq b_i, \quad i = 1, \dots, m$$

where the matrix  $\tilde{H}^{(i)} = \left[ (\Delta A_1)^T e_i^{(m)}, (\Delta A_2)^T e_i^{(m)}, \dots, (\Delta A_{|N|})^T e_i^{(m)} \right]$ .

*Proof.* To prove the result (i), we show that it is an immediate corollary of Theorem 6.2. To apply Theorem 6.2, we first reformulate the LP in the form (2.1) as we did at the end of Section 6.2. The  $i$ th constraint of  $Ax \leq b$ , i.e.,  $A_i x \leq b_i$  can be written as (6.15) where  $M^{(i)}$  is given by (6.14). Clearly, we have

$$\text{vec}(M^{(i)}) = e_i \otimes A_i^T, \quad \text{vec}(\bar{M}^{(i)}) = e_i \otimes \bar{A}_i^T.$$

Notice that when  $A_i$  belongs to the uncertainty set (6.16), then the  $\text{vec}(M^{(i)})$  belongs to the following uncertainty set:

$$\left\{ \text{vec}(M^{(i)}) \left| \exists u \in R^{|N^{(i)}|} : \text{vec}(M^{(i)}) = e_i \otimes \bar{A}_i^T + \sum_{j \in N^{(i)}} \left( e_i \otimes (\Delta A_j^{(i)})^T \right) u_j, \|u\|^{(i)} \leq \Omega^{(i)} \right. \right\}.$$

By Theorem 6.2, the robust LP is equivalent to

$$\begin{aligned} & \min c^T x \\ & \text{s.t. } \bar{a}_i^T x + \Omega^{(i)} \left\| \left( P^{(i)} \right)^T \chi_i \right\|_*^{(i)} \leq b_i, \quad i = 1, \dots, m, \\ & x \geq 0, \end{aligned}$$

where  $\chi_i = e_i \otimes x$  and the matrix

$$P^{(i)} = \left[ e_i \otimes (\Delta A_1^{(i)})^T, e_i \otimes (\Delta A_2^{(i)})^T, \dots, e_i \otimes (\Delta A_{|N^{(i)}|}^{(i)})^T \right].$$

Notice that

$$\left( P^{(i)} \right)^T \chi_i = \left[ (\Delta A_1^{(i)})^T, (\Delta A_2^{(i)})^T, \dots, (\Delta A_{|N^{(i)}|}^{(i)})^T \right]^T x.$$

Therefore, the result (i) holds.

Using the uncertainty set (6.18), item (ii) can also be proved by applying Theorem 6.2. In fact, we can reformulate the LP in the form of (2.1) as in Section 2 where all the data matrix  $M^{(i)}$  are equal to  $\begin{bmatrix} A \\ 0 \end{bmatrix}_{n \times n}$ . Notice that the uncertainty set (6.18) can be written as

$$\left\{ \text{vec} \left( \begin{bmatrix} A \\ 0 \end{bmatrix} \right) \mid \exists u \in R^{|N|} : \text{vec} \left( \begin{bmatrix} A \\ 0 \end{bmatrix} \right) = \text{vec} \left( \begin{bmatrix} \bar{A} \\ 0 \end{bmatrix} \right) + \sum_{j \in N} \text{vec} \left( \begin{bmatrix} \Delta A_j \\ 0 \end{bmatrix} \right) u_j, \|u\| \leq \Omega \right\}.$$

This is the uncertainty set of the form (6.4). Thus, by Theorem 6.2, robust LP is equivalent to

$$\begin{aligned} & \min c^T x \\ & s.t. \bar{a}_i^T x + \Omega \|H^T \chi_i\|_* \leq b_i, \quad i = 1, \dots, m, \\ & x \geq 0. \end{aligned}$$

where  $\chi_i = e_i \otimes x$  and the matrix

$$H = \left[ \text{vec} \left( \begin{bmatrix} \Delta A_1 \\ 0 \end{bmatrix} \right), \text{vec} \left( \begin{bmatrix} \Delta A_2 \\ 0 \end{bmatrix} \right), \dots, \text{vec} \left( \begin{bmatrix} \Delta A_{|N|} \\ 0 \end{bmatrix} \right) \right].$$

Denote by  $\tilde{\chi}_i = e_i^{(m)} \otimes x$  where  $e_i^{(m)}$  denote the  $i$ th column of the  $m \times m$  identity matrix. It is easy to check that

$$H^T \chi_i = \tilde{H}^T \tilde{\chi}_i = \left( \tilde{H}^{(i)} \right)^T x,$$

where the matrices

$$\tilde{H} = [\text{vec}(\Delta A_1), \text{vec}(\Delta A_2), \dots, \text{vec}(\Delta A_{|N|})],$$

$$H^{(i)} = [(\Delta A_1)^T e_i^{(m)}, (\Delta A_2)^T e_i^{(m)}, \dots, (\Delta A_{|N|})^T e_i^{(m)}].$$

Thus, the desired result (ii) follows.  $\square$

Notice that dual norms appear in (6.17) and (6.19). If the norms used are some special norms such as  $\ell_1, \ell_2, \ell_\infty, \ell_1 \cap \ell_\infty, \ell_2 \cap \ell_\infty$ , then their dual norms  $\|\cdot\|_*$  are explicitly known (see for example [14]).

In [12], Bertsimas, Pachamanova and Sim studied the case of robust LP with uncertainty sets defined by general norms. Their result provides a unified treatment of the approaches in [23, 24, 6, 7, 11]. However, their result is a special case of Theorem 6.7 above. Their uncertainty set is defined by the inequality

$$\|M(\text{vec}(A) - \text{vec}(\bar{A}))\| \leq \Delta.$$

where  $M$  is an invertible matrix and  $\Delta$  is a given constant. Clearly, this inequality can be written as

$$\text{vec}(A) = \text{vec}(\bar{A}) + M^{-1}u, \|u\| \leq \Delta.$$

This is a special case of the uncertainty model (6.18), corresponding to the case when  $|N|$  is equal to the number of data and the perturbation directions  $\Delta A_j$ 's are linearly

independent (here  $\Delta A_j$ 's are the column vectors of  $M^{-1}$ ). So, when we apply Theorem 6.7 (ii) to such a special uncertainty set, we obtain the same result as “Theorem 2” in [12]. But our result in Theorem 6.7 (ii) is more general than the result in [12] because our result can even deal with the cases when the perturbation direction matrix  $H$  is singular and even not a square matrix.

It should be mentioned that “Theorem 2” in [12] can also be obtained from our Corollary 6.6. Since  $M$  is invertible, we can define the function  $g(D) = \|MD\|$  which is a norm. The uncertainty set is defined by only one norm inequality, i.e.  $g(D) \leq \Delta$ . So, setting  $\ell = 1$  in Corollary 6.6, we obtain “Theorem 2” in [12] again.

Now we compare Theorem 6.7 with the corresponding results for robust LP in Bertsimas and Sim [14]. For LP, Theorem 6.7 (i) strengthens (generalizes) the corresponding result in [14] in the sense that we do not impose extra conditions on the norms, but in [14] a similar result is obtained under the additional assumption that the norms are absolute norms. Below we elaborate on this in more detail.

As we pointed out in Section 2, without loss of generality, it is sufficient to consider the case when only  $A$  is subject to uncertainty. For LP, only “row-wise” uncertainty is considered in [14]; for the  $i$ th linear inequality  $A_i x \leq b_i$ ,  $A_i$  belongs to the uncertainty set (6.16). Bertsimas and Sim [14] defined  $f(x, A_i) = -(A_i x - b_i)$ , and

$$s_j = g(x, \Delta A_j^{(i)}) =: \max\{-(\Delta A_j^{(i)})x, (\Delta A_j^{(i)})x\} = |(\Delta A_j^{(i)})x|, \quad j = 1, \dots, N^{(i)}.$$

Bertsimas and Sim [14] proved that for LP, when the norm  $\|\cdot\|^{(i)}$  used in (6.16) is an absolute norm, the robust LP constraint is equivalent to

$$f(x, \bar{A}_i) \geq \Omega^{(i)} \|s\|_*^{(i)} \quad (\text{or equally, } f(x, \bar{A}_i) \geq \Omega^{(i)} y, \quad \|s\|_*^{(i)} \leq y).$$

That is

$$-\bar{A}_i x - b_i \geq \Omega^{(i)} \left\| \left[ (\Delta A_1^{(i)})^T, (\Delta A_2^{(i)})^T, \dots, (\Delta A_{|N^{(i)}|}^{(i)})^T \right]^T x \right\|_*^{(i)}$$

which is

$$\bar{A}_i x + \Omega^{(i)} \left\| \left[ (\Delta A_1^{(i)})^T, (\Delta A_2^{(i)})^T, \dots, (\Delta A_{|N^{(i)}|}^{(i)})^T \right]^T x \right\|_*^{(i)} \leq b_i.$$

This is the same result as Theorem 6.7 (i). So, Bertsimas and Sim [14] proved the result of Theorem 6.7 (i) under the assumption that the norms used are absolute norms. We obtain this result without additional assumptions on the norms.

We can also apply our general results to nonlinear problems such as SOCP and QP. Let us comment on the differences of our approach from the approach of Bertsimas and Sim [14]. Applying our general results to robust QP would lead to *exact* formulations which, in general, would be computationally difficult. Bertsimas and Sim [14] aim at obtaining computationally tractable *approximate* formulations. These are two different ways of approaching nonlinear robust optimization problems. Computationally tractable approximate formulations are important for practical solution of large-scale problems: approximate solution is the price one has to pay for computational tractability. Exact formulations are also important. First, from theoretical viewpoint, they allow to gain more insight and to study the structure of the problems. Second, they can be used in practice to obtain exact solutions to small-scale problems. Third, they can provide new or strengthened results for important special cases when restricted to such cases, as demonstrated in this section.

**7. Conclusion.** One of our main goals was to show how the classic convex analysis tools can be used to study robust optimization. We showed that some rather general classes of robust optimization problems can be represented as explicit mathematical programming problems. We demonstrated how explicit reformulations of the robust counterpart of an uncertain optimization problem can be obtained, if the uncertainty set is defined by convex functions that fall in the space  $\mathcal{L}_H$  and satisfy the condition (4.1). Our strongest results correspond to the case where the functions defining the uncertainty set are homogeneous, because in this case the condition (4.1) holds trivially, and the robust counterpart can be further simplified. Our results provide a unified treatment of many situations that have been investigated in the literature. The analysis of this paper is applicable to much wider situations and more complicated uncertainty sets than those considered before; for example, it is applicable to cases where fluctuations of data may be asymmetric, and not defined by norms.

#### REFERENCES

- [1] I. Averbakh, On the complexity of a class of combinatorial optimization problems with uncertainty, *Math. Programming*, 90 (2001), pp.263-272.
- [2] I. Averbakh, Minmax regret solutions for minimax optimization problems with uncertainty, *Oper. Res. Letters*, 27 (2000), pp. 57-65.
- [3] I. Averbakh, Minmax regret linear resource allocation problems, *Oper. Res. Letters*, 32 (2004), pp. 174-180.
- [4] I. Averbakh and V. Lebedev, Interval data minmax regret network optimization problems, *Discrete Appl. Math.* 138 (2004), pp. 289-301.
- [5] A. Ben-Tal, *Conic and Robust Optimization*, Lecture Notes, University of Rome "La Sapienza", July 2002.
- [6] A. Ben-Tal and A. Nemirovski, Robust convex optimization, *Math. Oper. Res.*, 23 (1998), pp. 769-805.
- [7] A. Ben-Tal and A. Nemirovski, Robust solutions to uncertain linear programs, *Oper. Res. Letters*, 25 (1999), pp. 1-13.
- [8] A. Ben-Tal and A. Nemirovski, Robust solutions of linear programming problems contaminated with uncertain data, *Math. Programming*, 88 (2000), pp. 411-424.
- [9] A. Ben-Tal and A. Nemirovski, and C. Roos, Robust solutions of uncertain quadratic and conic-quadratic problems, *SIAM J. Optim.*, 13 (2002), pp. 535-560.
- [10] D. Bertsimas and M. Sim, Robust discrete optimization and network flows, *Math. Programming*, 98 (2003), pp. 49-71.
- [11] D. Bertsimas and M. Sim, The price of Robustness, *Oper. Res.*, 52 (2004), pp. 35-53.
- [12] D. Bertsimas, D. Pachamanova and M. Sim, Robust linear optimization under general norms, *Oper. Res. Letters*, 32 (2004), pp. 510-516.
- [13] D. Bertsimas and M. Sim, Robust discrete optimization under ellipsoidal uncertainty sets, Report. March, 2004, MIT
- [14] D. Bertsimas and M. Sim, Tractable approximations to robust conic optimization problems, *Math Program., Ser. B*, 107 (2006), pp. 5-36.
- [15] D. Bienstock, Experiments with robust optimization, Presented at *19th International Symposium on Mathematical Programming*, Rio-de-Janeiro, 2006.
- [16] Z. Chen and L.G. Epstein, Ambiguity, risk and asset returns in continuous time, Mimeo, 2000.
- [17] J. Dupačová, The minimax approach to stochastic program and illustrative application, *Stochastics*, 20 (1987), pp. 73-88.
- [18] J. Dupačová, Stochastic programming: minimax approach, In *Encyclopedia of Optimization*, Kluwer, 2001.
- [19] L.G. Epstein and M. Schneider, Recursive multiple priors, Technical Report 485, Rochester Center for Economic Research, August 2001.
- [20] L.G. Epstein and M. Schneider, Learning under ambiguity, Technical Report 497, Rochester Center for Economic Research, October 2002.
- [21] E. Erdoğan and G. Iyengar, Ambiguous chance constrained problems and robust optimization, *Math. Program., Ser. B*, 107 (2006), 37-61.
- [22] J. E. Falk, Exact solutions of inexact linear programs. *Oper. Res.*, 24 (1976), pp. 783-787.

- [23] L. El-Ghaoui and H. Lebret, Robust solutions to least-square problems with uncertain data matrices, *SIAM J. Matrix Anal. Appl.*, 18 (1997), pp. 1035-1064.
- [24] L. El-Ghaoui, F. Oustry and H. Lebret, Robust solutions to uncertain semi-definite programs, *SIAM J. Optim.*, 9 (1998), pp. 33-52.
- [25] D. Goldfarb and G. Iyengar, Robust convex quadratically constrained programs. *Math. Programming*, 97 (2003), Ser. B, pp. 495-515.
- [26] D. Goldfarb and G. Iyengar, Robust portfolio selection problems, *Math. Oper. Res.*, 28 (2003), pp. 1-38.
- [27] L.P. Hansen and T. J. Sargent, Robust control and model uncertainty, *American Economic Review*, 91 (2001), pp. 60-66.
- [28] R. Jagannathan, Minimax procedure for a class of linear programs under uncertainty, *Oper. Res.*, 25(1977), pp. 173-177.
- [29] P. Kouvelis and G. Yu, *Robust Discrete Optimization and Its Applications*, Kluwer Academic Publisher, Boston, 1997.
- [30] G. LeBlanc, Homogeneous programming: Saddlepoint and perturbation function, *Econometrica*, 45 (1977), pp. 729-736.
- [31] P.J. Maenhout, Robust portfolio rules and asset pricing, *The Review of Financial Studies*, 17 (2004), pp. 951-983.
- [32] P.V. Moeseke. Saddlepoint in homogeneous programming without Slater condition, *Econometrica*, 42, No.3 (1974), pp.593-596.
- [33] J.M. Mulvey, R. J. Vanderbei and S.A. Zenious, Robust optimization of large scale systems, *Operations Research*, 43 (1995), pp. 264-281.
- [34] J.M. Ortega, and W.C. Rheinboldt, *Iterative Algorithms for Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
- [35] M.C. Pinar and R.H. Tütüncü, Robust profit opportunities in risky financial portfolios, *Oper. Res. Lett.* 33 (2005), pp. 331-340.
- [36] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, Second Edition, 1970.
- [37] A. Shapiro and S. Ahmed, On a class of minimax stochastic problems, *SIAM J. Optim.*, 14 (2004), 1237-1249.
- [38] A. Shapiro and A.J. Kleywegt, Minimax analysis of stochastic problems, *Optim. Meth. & Software*, 17 (2002), pp. 523-542.
- [39] A. L. Soyster, A duality theory for convex programming with set-inclusive constraints, *Oper. Res.*, 22 (1974), pp. 892-898.
- [40] A. L. Soyster, Convex programming with set-inclusive constraints and applications to inexact linear programming, *Oper. Res.*, 21 (1973), pp. 1154-1157.
- [41] Z. Wang, A shrinkage approach to model uncertainty and asset allocation, *The Review of Financial Studies*, 18 (2005), 673-705.
- [42] J. Žáčková, On minimax solutions of stochastic linear programs, *Čas. Pěst. Mat.*, 91 (1966), pp. 423-430.
- [43] Y. Zhang, A general robust optimization formulation for nonlinear programming, Department of Computational and Applied Mathematics, Rice University, June 2005.